DOCUMENT RESUME

ED 127 050 EC 009 298

AUTHOR Hiat, Albert B.; And Others
TITLE Evaluation of the Albuquerque Indian School Motivational Environment Program.
INSTITUTION Albuquerque Indian School, N. Mex.; Bureau of Indian Affairs (Dept. of Interior), Albuquerque, N. Mex.
PUB DATE Aug 71
NOTE 180p.

EDRS PRICE MF-$0.83 HC-$10.03 Plus Postage.
DESCRIPTORS Academic Achievement; *American Indians; Attendance; *Behavior Change; Drinking; *Motivation Techniques; *Program Evaluation; *Secondary Education; Self Concept; Student Attitudes; Student Behavior; Teacher Attitudes
IDENTIFIERS *Albuquerque New Mexico Indian School; *Token Economy

ABSTRACT

In order to evaluate the impact of a token economy behavior modification program implemented from 1970-71 in the Albuquerque Indian School (AIS), a secondary institution, a five-member evaluation team assessed standardized test results, behavioral data, and student and staff attitudes. A battery of tests (Tennessee Self-Concept Scale, SRA Achievement Series, and/or the Illinois Test of Educational Development) were administered to 41 students at AIS on a pre- and post-test basis. Employing the reversal technique, behavioral data taken from the dormitories and classrooms were gathered on drinking, absenteeism without leave, tardiness (bedcheck), and class absence and tardiness. Personal interviews (N=63) and open-ended questionnaires were utilized to assess student attitudes toward the school in general, the token economy, perceptions of change, future programs, etc. Staff attitudes were evaluated in terms of a conceptual systems theory. Evaluation findings suggested the following: the most important conclusions were based on subjective data; most of the objective data indicated the program had not significantly altered behavior; the behaviors chosen for modification were extremely complex, deeply rooted, and influenced by external factors; lack of staff commitment, training, and/or opportunity contributed to project problems; some individual improvement was apparent. (JC)

****
* Documents acquired by ERIC include many informal unpublished *
* materials not available from other sources. ERIC makes every effort *
* to obtain the best copy available. Nevertheless, items of marginal *
* reproducibility are often encountered and this affects the quality *
* of the microfiche and hardcopy reproductions ERIC makes available *
* via the ERIC Document Reproduction Service (EDRS). EDRS is not *
* responsible for the quality of the original document. Reproductions *
* supplied by EDRS are the best that can be made from the original. *
****

ED127050

U S DEPARTMENT OF HEALTH
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

EVALUATION

ALBUQUERQUE INDIAN SCHOOL

MOTIVATIONAL ENVIRONMENT PROGRAM

2 August 1971

PS009298

EVALUATION OF THE ALBUQUERQUE INDIAN SCHOOL

MOTIVATIONAL ENVIRONMENT PROGRAM

Albert B. Hiat
Evaluation Project Director
University of New Mexico
College of Education
Department of Educational Foundations
Albuquerque, New Mexico 87106

August 1971

The research reported herein was performed pursuant to a contract with the Bureau of Indian Affairs, Albuquerque Indian School, Department of Interior. Contractors undertaking such projects under Government sponsorship are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official Bureau of Indian Affairs position or policy.

U.S. Department of Interior
Bureau of Indian Affairs

Albuquerque Indian School

Preface

The purpose of the project reported in this document was to evaluate and to provide input into the Motivational Environment Program that was in operation at the Albuquerque Indian School for the 1970-1971 school year.

The members of the evaluation team were not employees of the AIS but were an independent group working through the University of New Mexico, Department of Educational Foundations. The Motivational Environment Program had been operational for most of the year before the present evaluation was initiated. Testing instruments had already been selected by the AIS staff and pre-testing had been done early in the school year. The format and method of collecting the behavioral data were also decided upon before the evaluation team appeared on the scene.

The method of collecting and analyzing the staff and student attitudinal and interview data was decided upon by the evaluation team which selected and developed all the necessary instruments. Because of this split in planning and responsibility and the late date at which the evaluation team began its task, some problems were encountered in acquiring and interpreting data already collected, such as the pre-test data and some of the behavioral data.

It is strongly recommended, if an evaluation of this sort is attempted again, that an evaluation team be selected before the program begins so that they can collaborate in the formulation of the evaluation design and development and be responsible for the data collection system.

An evaluation of this kind of program would also be more meaningful and conclusive if a control group were utilized. This would appear to be necessary because of all of the uncontrolled variables that are in operation in an environment such as the AIS where the students are exposed to many activities and people who may affect their behavior.

Albert B. Hiat
Project Director

Acknowledgements

We wish to express our appreciation to Dr. Albert W. Vogel, Dr. James Cooper and Dr. Richard Harris for their invaluable contributions to the evaluation project.

We also wish to thank the following undergraduate and graduate students for all the difficult and often tedious tasks that they performed so efficiently:

Mike Allison
Kathleen Aragon
Alvin Begay
Marlene Benally
Anna Boyd
Ophelia Charlie
Ralph Charlie
Conroy Chino
Diana Dempsey
Paul Dion
Jonathan Embry
Connie Garcia
Mary Gatewood
Jo-Ann Godfrey
Alice Hazen
Melvin Kanner
Jon G. Knudsen
Richard Kyrlach
Alice Luna
Sonya McKinney
Gary Mills
Thomas Nahkai
Deborah Neuman
James Nez
Phil R. Pohl
Carla Scott
Deborah Tharinger
Robert Tohe
David VanMeveren
Lee Ann VanMeveren
Ryan Wagner
Genevieve Yazzie

Special thanks are in order for Jean Berch and Molly Attruia for their patience and the speed and accuracy of their typing and the excellent job they did assembling the final manuscript.

Table of Contents

		Page
SECTION I:	Rationale for the use of a token economy at the Albuquerque Indian School. Joseph D. Blanchard, Psychologist, Albuquerque Indian School.	1
SECTION II:	Analysis of standardized test results. Alice Biernoff Hiat, Ph.D.	4
	References	21
	Appendix A	22
SECTION III:	Analysis of behavioral data. Daniel B. Berch, Ph.D.	31
	References	61
SECTION IV:	Assessment of student attitudes and perceptions. Marcia Landau, M.S.	62
	Appendix B	97
SECTION V:	Assessment of staff attitudes and perceptions. Wayne P. Moellenberg, Ph.D.	104
	References	142
	Appendix C	143
SECTION VI:	Final conclusions and recommendations. Project Evaluation Team.	163

Section I

Rationale for the Use of a Token Economy at the Albuquerque Indian School

by

Joseph D. Blanchard, Psychologist
Albuquerque Indian School

Rationale for the Use of a Token Economy

at the Albuquerque Indian School

The MEP has been conceived as a means of instituting change in what heretofore had been perceived a rigid, closed-system institution. Its primary purpose, at least inferentially, has been to break up the hard crust of the old process so that change might be instituted with some hope of future success. One might argue (as has been done) that to attack the total school system head-on is a high risk process destined to result in failure. Indeed, some have stated that such an approach is foolhardy. Those who speak such words often talk of "pilot" studies, of taking a class or some other small part of the system and proving your approach. I do not deny that such approaches would be more "safe"; be neater; would have less risk and a higher probability of "experimental" success. All of these considerations were examined beforehand.

The decision to go whole-hog, to address (even attack) the total school process was a committment to certain values. The historical ground of the BIA is littered with good programs which fell to the wayside. Most of these programs represented small-scale, safe efforts which had no impact on the total system process. Most were smoothed over much as the oyster coats a grain of sand, to be shoved aside and soon forgotten. The committment made and represented by the MEP is to cause permanent structural change in the AIS system process in full knowledge that such a high risk approach might polarize resistance, perhaps even closing the system tighter, thus making it more impermeable. However, these were two major factors mitigating against that possibility: (1) a new and young administration; (2) many of the teaching staff being eligible for retirement. Therefore, the decision was

made to confront the entire system, realizing that much time and effort (and compassion) would be required to handle the resistance, anxiety, and fear.

A behavioral modification program was chosen as the most viable means of propping up the system while the tearing down and restructuring of processes was occurring. The token economy in the overall system process has been considered all along as an emergency intervention device. Prosthetics are emergency supports, never any substitute for the original, the natural, the intrinsic. The same holds true for extrinsic motivational prosthetics. In the long run there is no substitute for intrinsic motivation, whatever the process. We hold that to be true in this case.

As the evaluation report indicates, there was nothing neat about the MEP at AIS. There was no attempt made at control, at a convincing experimental design. Objective evidence indicates the program failed miserably in meeting some of the objectives, in securing some of the desired behavioral change. But change has occurred in the system. Most staff members have survived the onslaught and are going on to bigger and better programs. Some who chose not to have left. Many have retired. Innovative programs are being started. A new token economy is being instituted incorporating many of the recommended changes. Change is in the air at AIS.

Has it been worth the effort? I think so! Who can measure all the ramifications of such an effort? I think no one.

I have been all along as much concerned with values and feelings as with facts and concepts. The ultimate concern among people is not what they know but what their committments are. Thus, the final distinction among people pertains to their values and feelings, not their knowledge and skills. What we have needed at AIS has been committment to the children and to

approaches having some probability of success in child development and achievement. You have to know you are lost before you can make efforts to find yourself. The MEP was designed to demonstrate with impact how much in error the system has been.

Salvation has always been dependent on recognition, repentence, restitution and committment to change. The salvation of AIS as an educational institution of worth is no less dependent on these factors.

Remember! We are not entirely in the realm of facts. We are not in a situation demanding a problem-solving mode. We are in a human condition requiring the best in value judgements. What we require is not more knowledge but committment to what we already know. Through the MEP such a committment has been put into action. The situation has demanded it. The development and education of these Indian children have made such a committment imperative.

Section II

Analysis of Standardized Test Results

Alice Biernoff Hiat, Ph.D.

Analysis of Standardized Test Results

A battery of tests including the Tennessee Self-Concept Scale, and the SRA Achievement Series or the Illinois Test of Educational Development (depending on grade in school) was administered to the students at the Albuquerque Indian School in May 1970 and again in May 1971 as part of the project evaluating the impact of the Motivational Environment Program token economy instituted in October 1970. The design of the study, the choice of tests used, and the particular scores reported were decided upon by the AIS staff and the actual testing of the students was carried out under the supervison of the AIS school personnel prior to the writer's joining the evaluation team.

Tennessee Self-Concept Scale (TSC). The rationale for assessing changes in self-concept of the AIS Students from pre- to post-testing was based on the assumption that self-concept reflects the individual's experiences in the "real world." It was presumed that if the Motivational Environment Program were successful in changing the behavior of the students involved in it in a positive direction, these changed behaviors would provide more success experiences for the students and, in turn, would be reflected in improved self-concept. There is a vast literature on the relationship between self-concept and school achievement as well as self-concept and functioning in general (for a review of the relevant literature, see McCandless, 1967, Ch. 6), in light of which it seemed reasonable to assume that improved functioning would be associated with more positive self-perception.

The TSC was developed to provide a standardized multi-dimensional scale for the measurement of self-concept. The scale consists of 100 self-descriptive statements which the subject uses to present his own picture of himself. The author, William H. Fitts, states in the TSC manual (1965) that the scale is self-administering for either individuals or groups and can be used with subjects age twelve or older having at least a sixth grade reading level.

The scale was standardized on a sample of 626 persons from various parts of the country, ranging in age from 12 to 68. Included were approximately equal numbers of males and females, both black and white subjects, representative of all social, economic, intellectual, and educational levels from sixth grade through the Ph.D. degree. Fitts (1965) states that the effects of such demographic variables as sex, age, race, education, and intelligence on the scores of the scale are quite negligible. (However, the only data for race referred to in the manual is some for black nursing students and black college students, both of which groups seem to be quite similar in many respects to the college students whom Fitts admits are over-represented in the norm group.)

The validity of using the TSC as a measure of the self-concept of the students at the Albuquerque Indian School would appear to be questionable on a number of counts. First, the TSC requires reading ability equal to that of the sixth grade level. The results of achievement testing on the AIS students suggested that the reading level of many of the ninth to twelfth grade students (to whom the scale was administered) may have been below this. Second, despite the assurance of the author to the contrary, the present writer feels that the instructions and procedures for taking the test are somewhat complicated and would require close supervision of the subjects. For example, the items are not numbered consecutively, the subjects must skip lines on the answer sheets when progressing from item to item, and the columns

in which answers are recorded go from right to left rather than the more familiar left to right. It is felt that these features of the TSC would provide some degree of distraction for subjects taking the test, and this in conjunction with the poor reading ability of the AIS students raises some question regarding the validity of the TSC results on this group.

Some support for the assumption that the students were able to perform the mechanics of the test adequately is seen in the fact that the pre- and post-test profiles are quite similar (see Fig. 1), and in the fact that the subtest standard deviations for the AIS group were fairly close to the standard deviations reported for the norm group. This latter implies that the AIS students were no more variable in their responses to the various scales than was the norm group; one would expect more variability if the students had been responding more or less randomly because of inability to follow testing procedures. The main suspicious feature of the data is the high T/F ratio (see below) which indicates that as a group the AIS students tended to respond in a positive direction to statements no matter what their content. This could have occurred if the students had not been trying to respond appropriately to the items but merely tending to mark items "true of me" down the columns of the answer sheet.

Third, despite Fitts' insistance that "race" does not affect scale scores, it is felt that the adolescent Indian population of AIS is probably quite different in many ways from the black nursing and college students referred to as examples of "race" in the TSC manual. It would seem that Fitts' use of the term "race" refers primarily to skin color and not to what might better be termed cultural differences. It seems quite likely that the items pertaining to the TSC scoring categories of "physical self" (where the individual presents his view of his body, state of health, physical appearance, skills

and sexuality); "moral ethical self" (where the individual presents his view of his moral worth, relationship to God, feelings of being a good or bad person, and satisfaction with his religion); "personal self" (reflecting the individual's sense of personal worth); "family self" (reflecting the individual's feelings of adequacy, worth, and value as a family member); and "social self" (reflecting the individual's sense of personal adequacy and worth in his interactions with other people in general) might have quite different meanings to persons with a culture as different from that of the middle-class college students as that of the various Indian cultural groups. There is no way of determining to what extent this alteration of meaning in test items might have occurred in the present situation. However, the possibility that this did occur should be kept in mind when interpreting the results of the testing.

Results. There were 41 students, including 25 females and 16 males, ranging in age from 14 to 20 and in grade from ninth to twelfth, for whom data were provided for both pre- and post-testing on the TSC. Means and standard deviations for the scale scores provided are presented in Table 1.

"Total P," reflecting the overall level of self-esteem, is stated by Fitts (1965) to be the most important single score in the TSC. For this reason a t-test for correlated means was performed to determine if the increase of 7.10 in Total P was significant; it was found to be so (t = 2.22, significant at the .05 level). An examination of individual scores revealed that 24 subjects showed an increase in Total P from pre- to post-testing, 15 showed a decrease, and two showed no change.

A multiple regression was done with change in Total P against sex, grade, and age to determine if any of these variables were associated with the change scores for this scale. Results of a stepwise regression analysis indicated that change in Total P was significantly associated with sex (multiple r = .315);

Table 1

Tennessee Self-Concept Scale:
Subtest Means and Standard Deviations
for Pre- and Post-Testing (N = 41)

	Pre-Test		Post-Test		Difference
	Mean	S.D.	Mean	S.D.	
Total P	309.78	(28.75)	316.90	(29.63)	7.12*
P Score: Identity	111.15	(13.05)	113.39	(12.49)	2.24
P Score: Self Satisfaction	96.05	(13.35)	98.00	(11.88)	1.95
P Score: Behavior	102.59	(9.23)	105.51	(11.09)	2.92
P Score: Physical Self	65.24	(7.73)	66.78	(6.51)	1.54
P Score: Moral-Ethical Self	60.29	(6.52)	59.78	(6.61)	-0.51
P Score: Personal Self	60.73	(7.15)	63.24	(7.08)	2.51
P Score: Family Self	64.46	(8.12)	66.51	(8.10)	2.05
P Score: Social Self	59.05	(7.06)	60.59	(8.54)	1.54
Self-Criticism	31.63	(4.18)	30.27	(5.29)	-1.36
Variability	45.68	(14.04)	44.27	(10.91)	-1.41
Distribution	94.83	(28.33)	101.41	(30.63)	6.58
True-False Ratio	1.37	(0.64)	1.21	(0.60)	-0.16
Net Conflict	10.66	(18.49)	4.80	(23.87)	-5.86
Total Conflict	39.59	(12.89)	38.95	(14.18)	-0.64
Defensive Positive	56.85	(11.01)	58.90	(11.56)	2.05
General Maladjustment	82.93	(8.61)	87.54	(8.79)	4.61**
Psychosis	58.93	(7.05)	57.29	(6.57)	-1.64
Personality Disorder	61.71	(9.30)	64.58	(10.08)	2.87
Neurosis	78.12	(8.40)	79.88	(10.02)	1.76
Personality Integration	7.41	(3.46)	7.46	(3.58)	0.05

* significant at .05
**significant at .001

and with grade (the addition of this variable brought the multiple r to .410). F-ratios for these r's were 4.28 and 3.92 respectively, both significant at the .05 level.

These analyses indicated that increase in Total P was associated with being female, and with being in a lower grade. Mean change in Total P for males was -.90 (7 subjects increased, 7 decreased, 2 showed no change); mean change for females was +12.24 (17 increased, 8 decreased). Breaking the data down by grades revealed that the mean change was +14.44 for the ninth grade

(6 increased, 3 decreased); +14.00 for the tenth grade (11 increased, 6 decreased); +4.00 for the eleventh grade (5 increased, 2 decreased, 2 showed no change); and -2.33 for the twelfth grade (2 increased, 4 decreased). The behavioral data of subjects having large increases and those having large decreases in Total P were examined to determine if changes in Total P was associated with behavioral changes; no systematic relationship was found between these variables. (See Berch chapter below on analysis of behavioral data.)

A *t*-test for correlated means was performed to determine if the increase in the mean GM (general maladjustment) score (which implies decrease in maladjustment), was significant; it was found to be so ($t = 3.49$, significant at the .001 level). Examination of the data for individuals showed that 30 increased, 10 decreased, and one showed no change in this score.

Discussion of TSC Results. An inspection of the profile for the mean subject scores for the AIS students, pre- and post-testing, reveals some interesting features (see Fig. 1). For persons not familiar with the TSC, a description of the scales is included in the Appendix.

The mean True/False ratio was high and deviant. This score reflects response set, in this case a tendency beyond normal limits to answer items in a positive direction regardless of content. This may have been a result of test-taking difficulties (see discussion above), or, to quote Fitts (1965), "It can be treated purely as a task approach or behavioral measure which has meaning only in terms of empirical validity. In this sense the T/F ratio differentiates patients from nonpatients and correlates significantly with other tests," (e.g., the *F*, Hysteria, Schizophrenia, Hypomania, and Paranoia scales on the MMPI, which are associated with bizarre responses, emotional immaturity, lack of control over behavior, and suspiciousness and hostility).

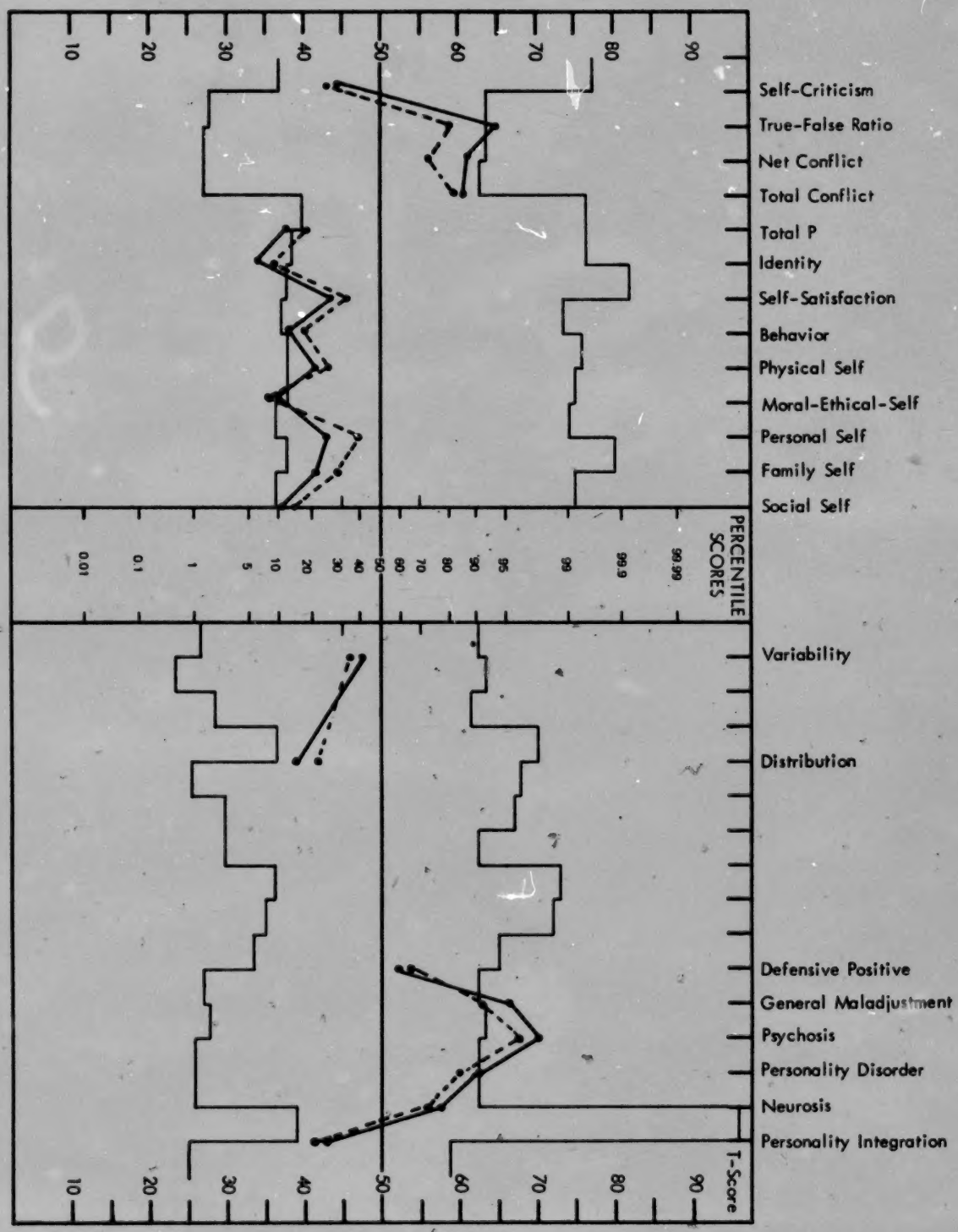

Fig. 1. TSC profiles of AIS students, *pre-* and *post-testing*

(See Wolman, 1965, p. 461, for a discussion of empirically based personality tests.) This latter interpretation is consistent with the high and deviant AIS group score on the Psychosis scale, as well as on (at least at pre-testing), the General Maladjustment and Personality Disorder scales. These scales are all empirical, i.e., they are composed of items that have been found to discriminate these various diagnostic groups from one another, rather than of items that have been chosen because they appear to experienced judges to be relevant to the personality characteristic under consideration, as is the case with other scales in the TSC. Fitts (1965) states that a high T/F ratio suggests a person with ". . . a weak ego and poor controls over . . . own behavior, likely to act out . . . conflicts and to be easily influenced by others (p. 13)." Thus, it would appear that the AIS students responded to the items on the TSC in a manner similar to that in which psychiatric patients in the Anglo society respond. This does not necessarily imply that the AIS students as a group are severely emotionally disturbed, but may reflect cultural differences in the meanings of various items. This finding certainly bears further study, both psychologically and anthropologically.

Another interesting feature of the group profile is the fact that the Self-Satisfaction score, though below the 50th percentile for the standardization group, is a great deal higher than the Identity and Behavior scores, which were both close to the lower limits of normality. This implies that while as a group the AIS students tend to report very low self-evaluation on basic identity and perceived actual behavior, they are relatively comfortable about this low level of functioning. This response to perceived inadequacy has somewhat negative implications for motivation to improve, at least in the context of Anglo values. (See Finney, 1969, pp. 159-167, for an interesting discussion of the validity of interpretations of psychological tests, in this

case the Rorschach, administered to persons of varying cultural background.)

Summary. It would appear that as a group the AIS students, on both pre- and post-testing, responded to items on the empirical scales much as do Anglo psychiatric patients who show bizarre behavior, have a weak ego, poor controls over their behavior, and a tendency to act out conflicts and to be easily influenced by others. The nonempirical, "Positive Score" scales were all below the 50th percentile of the standardization group, but, except for Identity, were within normal limits.

On post-testing, there was a significant increase in the Total P score, mainly on the part of girls in the ninth and tenth grades, implying that these students underwent some improvement in their self-concept, in the sense that they attributed more positive statements to themselves on post-testing than on pre-testing. (However, this "improvement" was not associated with systematic changes in the behavioral measures, as discussed in the following chapter by Berch.) This interpretation depends on the assumption that these students were able to respond appropriately to the testing situation, in that they understood directions, etc. The fact that the group profiles were similar for both pre- and post-testing tends to support this assumption, as does the fact that the standard deviations for the AIS group on the various scales were quite similar to those of the norm group.

However, the design of the present study does not allow one to make any causal statements about the reason for this significant increase in Total P. In order to attribute credit for this apparent improvement to the Motivational Environment Program, it would be necessary to have data on students at AIS not involved in the program but similar to the program students in all other respects. This was not done in the present study, and thus no conclusive statements can be made as to the effectiveness of the program in the area of

improvement of self-concept.

The fact that the increase in Total P in the present study occurred primarily in a specific group, i.e., ninth and tenth grade girls, raises the question of whether there was some particular factor operating within the total program having nothing to do with the token economy, e.g., one staff person who was particularly supportive to the girls in question. Here again, the necessity of having a control group in order to interpret differences from pre- to post-testing is evident.

SRA Achievement Series - Illinois Test of Educational Development (ITED). The severity of the educational retardation of the students at the Albuquerque Indian School has been pointed out and discussed by a number of writers (Moeny, 1971; Reedy, undated; AIS FY 1971-1972 Title I Proposal, undated).

Reedy (undated) found that the performance of sixth to ninth grade students on the SRA Achievement Series ranged from one-and-a-half to three years below actual grade level, and that the performance of tenth to twelfth graders on the ITED was at the first to fifth percentile, with percentile standing decreasing from the tenth to the twelfth grade.

Moeny (1971) reported a study in which the Wide Range Achievement Test (WRAT) was administered to seventh to ninth grade AIS students in Spring 1971. It was found that, for the reading subsection, the seventh and eighth graders had a mean grade level of 5.4 and the ninth graders a mean grade level of 5.8. For the arithmetic subsection, the seventh and eighth graders had a mean level of 4.05, and the ninth graders had a mean grade level of approximately 4.65.

This extremely low level of academic achievement is considered by the staff of the Motivational Environment Program to be the outcome of a combination of English language deficiency, deficiency in self-concept and social relationships and alienation, among other factors. As such, it was anticipated

that changes for the better in these areas as a result of the Motivational Environment Program would be reflected in improved functioning in the academic area. In order to evaluate any such changes, the SRA and the ITED were administered prior to and following the period during which the token economy was in effect.

SRA results. Data were provided for both pre- and post-testing for 32 seventh to ninth graders, including 19 females and 13 males. Data on age were not provided. Means and standard deviations for the SRA subtests are presented in Table 2.

Table 2

SRA Achievement Series: Subtest Means and Standard Deviations for Pre- and Post-Testing Seventh to Ninth Grade Students (N = 32)

		Pre-Test Percentile		Post-Test Percentile		Difference
		Mean	S.D.	Mean	S.D.	
	Social Studies	17.16	(19.83)	11.72	(11.97)	- 5.44
	Science	15.25	(18.44)	11.13	(10.62)	- 4.12
Language Arts	Capitalization/ Punctuation	21.75	(17.73)	10.09	(9.33)	-11.66
	Grammar	19.87	(16.57)	14.31	(13.30)	- 5.56
	Spelling	28.53	(23.09)	24.16	(22.21)	- 4.37
	Total Language Arts	20.59	(17.93)	13.03	(12.39)	- 7.56
Arithmetic	Reasoning	11.46	(10.94)	7.44	(7.45)	- 4.02
	Concepts	8.53	(11.98)	5.88	(6.11)	- 2.65
	Computation	17.71	(18.91)	6.72	(8.96)	-10.99
	Total Arithmetic	9.13	(9.95)	4.56	(6.14)	- 4.57
Reading	Comprehension	10.53	(11.08)	8.09	(8.97)	- 2.44
	Vocabulary	10.87	(10.95)	7.50	(10.00)	- 3.37
	Total Reading	9.06	(10.21)	6.75	(8.24)	- 2.31
	Composite	10.28	(11.89)	5.97	(6.90)	- 4.31

It should be noted that in most cases the figures for the means were smaller than those for the standard deviations. This reflects a skewed distribution and indicates that the group means were inflated by a few high-

scoring students. This implies that the level of achievement for the group as a whole was even lower than it appears, i.e., most students performed even more poorly than the means suggest.

A t-test for correlated means was performed to determine if the decrease of 4.31 in the Composite percentile score was significant; the t of 3.19 was significnat at the .01 level. Profiles of subtest mean scores for pre- and post-testing are presented in Figure 2.

A canonical correlation was done to determine if the change in individual percentile scores from pre- to post-testing for all subtests was related to sex and grade. A canonical r of .87 was obtained, which was not significant at the .05 level, indicating that individual changes in subtest percentile scores were not significantly associated with sex or grade.[1]

ITED results. Data were provided for both pre- and post-testing for 43 eleventh and twelfth graders (no complete data were provided for tenth graders), including 23 females and 20 males. No data were provided for age.

Means and standard deviations for the ITED subtests are presented in Table 3.

Table 3

Iowa Test of Educational Development (ITED): Subtest Meand and Standard Deviations for Pre- and Post-Testing, Eleventh and Twelfth Grade Students (N = 43)

Subtest	Pre-Test Percentile		Post-Test Percentile		Difference
	Mean	S.D.	Mean	S.D.	
1. Social Studies Concepts	7.63	(2.98)	8.05	(2.99)	0.42
2. Natural Science	6.05	(3.14)	7.60	(3.87)	1.55
3. Expression	10.30	(3.05)	9.53	(3.56)	-0.77
4. Quantitative Thinking	9.63	(3.70)	6.88	(2.42)	-2.75
5. Reading Social Studies	9.51	(2.76)	8.42	(3.05)	-1.09
6. Reading Natural Science	7.26	(3.42)	7.33	(3.12)	0.07
7. Interpreting Literary Materials	8.77	(2.33)	8.56	(2.40)	-0.21
Average of 5, 6, 7	9.51	(3.13)	9.98	(2.93)	0.47
8. General Vocabulary	6.14	(3.63)	5.93	(3.67)	-0.21
Composite: 1-8	7.33	(2.55)	7.09	(2.79)	-0.24
9. Use of Sources of Information	8.58	(6.07)	8.00	(4.27)	-0.58

[1]The chance level for the canonical correlation is considerably greater than zero. (see Morrison, Ch. 6).

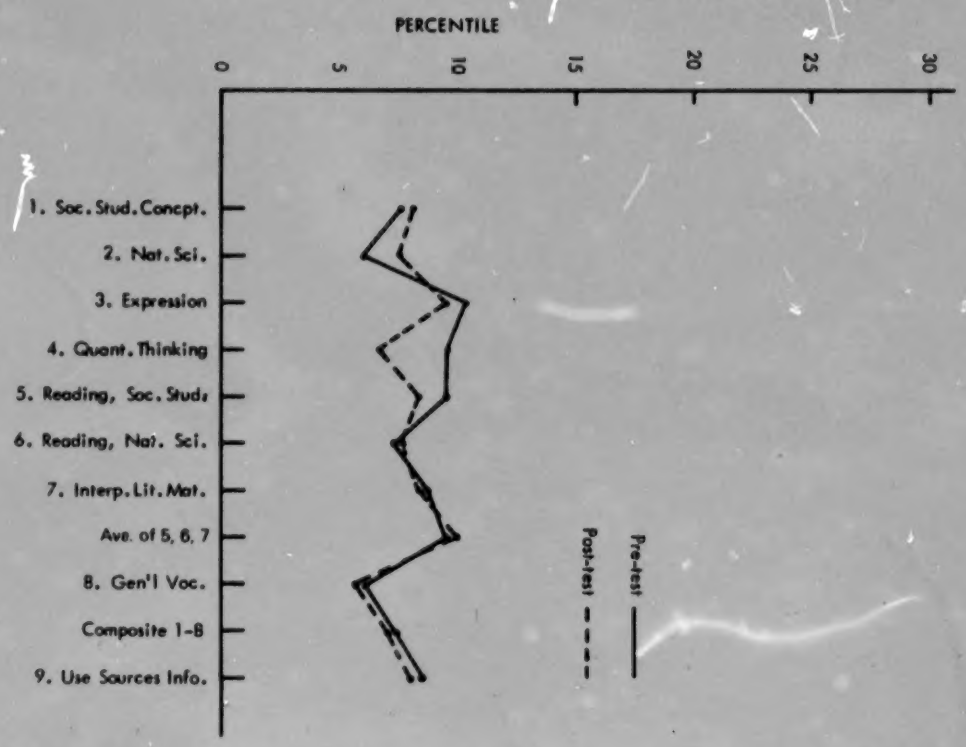

Fig. 3. ITED subtest profiles of eleventh and twelfth Grade AIS students, *pre* and *post-testing* (N=43).

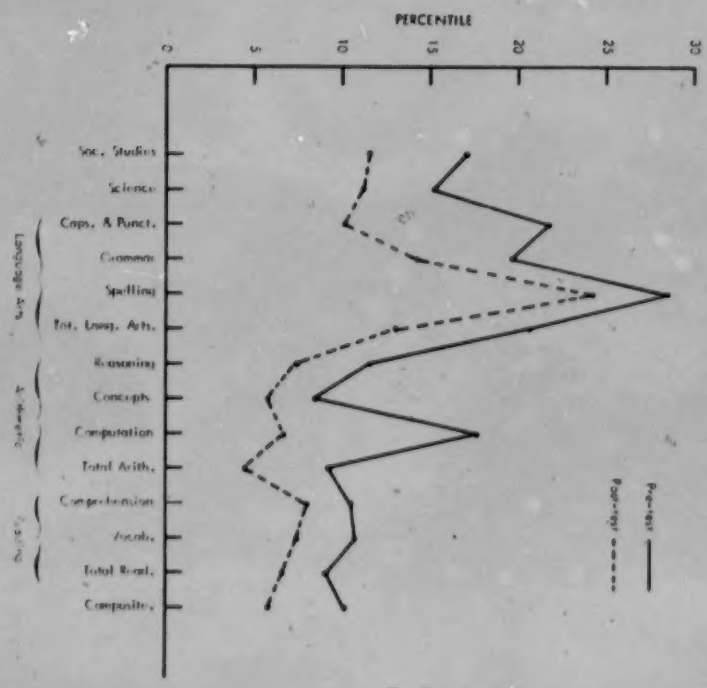

Fig. 2. SRA subtest profiles of seventh to ninth grade AIS students, *pre-* and *post-testing* (N=32).

A *t*-test for correlated means was performed to determine if the decrease of -0.24 in the composite score was significant; it was not (t = 1.32).

Profiles of subtest mean scores for pre- and post-testing are presented in Figure 3.

A canonical correlation was done to determine if the change in individual percentile scores from pre- to post-testing for all subtests was related to sex and grade. A canonical r of .51 was obtained, which was not significant at the .05 level, indicating that changes in individual subtest percentile scores were not significantly associated with sex or grade.

Discussion of Achievement Test Results. The academic performance of the AIS students is abysmally poor as measured by these standardized achievement tests, and indicates that as a group these students are ill-prepared to function in any situation where academic skills are required. The present evidence suggests that the AIS students actually deteriorate in level of performance relative to the norm group as they go from the sixth - ninth grade (SRA) to the eleventh - twelfth grade (ITED), as seen by comparison of the profiles for these two groups (Figs. 2 and 3). However, it should be noted that part of this apparent extreme drop is due to the fact that the SRA scores were inflated by a few extremly high-scoring subjects; thus, the ITED performance was not that much higher but rather the SRA means give an inflated impression of the actual level of performance of most of that group.

Nevertheless, the significant decrease in level of performance from pre- to post-testing for the SRA group does suggest that there is a true deterioration in academic performance relative to the norm group as one ascends the grade levels at AIS. This apparent deterioration may perhaps more accurately be described as a failure to gain at the same rate, if at all, as the norm group, which is then evidenced as a decline in standing relative to the norm

group. This finding is consistent with that of Reedy's (undated) who also found a decrease in mean percentile score on the ITED from tenth to twelfth grades for AIS students.

In view of the present SRA data showing a significant drop from pre- to post-testing, and of Reedy's data, it is of interest to note that for the present group of ITED subjects there was not a significant decrease in the composite score in the inter-test interval. If the expectation is that these subjects should have steadily decreased in percentile score, then the absence of a decrease represents "improvement" (although all the post-test means are still below the 10th percentile, which makes the use of the term "improvement" in this context a strictly technical one). However, there is no way to relate this apparent "improvement" in scholastic achievement to the intervening implementation of the Motivational Environment Program, as there was no data for a control group of equivalent subjects not exposed to this program.

It should be noted additionally that despite the fact that the ITED mean percentile scores were quite low, they were higher (all above the 6th percentile) than were the scores in Reedy's data (all below the 4th percentile). This discrepancy may be related to the fact that the present data includes only those subjects who had both pre- and post-test data, i.e., those who were in attendance at the Indian School from August 1970 to May 1971. It is possible that this group represents the "better" students, the poorer ones having dropped out of school (and thus out of the study) in the interim.

Recommendations. It is recognized that the difficulties involved in implementing an effectively controlled evaluation study in a school setting are multitudinous. However, the importance of control group data to a meaningful interpretation of changes that may occur subsequent to the initiation of a new program is such that every attempt should be made to ensure that a

control group is incorporated into any evaluation plan, that the control group is free of exposure to the program being studied, and that the pre- and post-assessment of the control group is carried out as carefully as that on the group exposed to the program. Only when this is done can one make conclusions about the impact of a program with any degree of confidence in one's assertions. In the present study, it might have been possible to implement the token economy only in certain classes, or only in certain dorms; failing this, it might have been possible to collect control data from another Indian School with a similar population, although this alternative would have been less satisfactory for a number of reasons. In any case, the expense and work involved in providing good control group data is well repaid by the increase in the certainty with which one can state conclusions about the effectiveness of a given program.

A greater effort should be made in any program evaluation using standardized tests to ensure that complete data is obtained from as many subjects as possible. In the present evaluation using only subjects with pre- and post-testing, data on many subjects had to be discarded because it was incomplete, i.e., either pre- or post-test data were missing. It is not possible to post-test subjects who have dropped out of the study, but it is possible to avoid spending the time and effort post-testing students who have not been pre-tested. Also, once the investment of time and effort has been made in pre-testing a student, every effort should be made to obtain post-test data from him if he is still available. Another kind of problem with incomplete data occurred in the present study because many students who did have complete pre- and post-test data on one test did not have it on another. For example, it had been planned to include the change score for Total P on the TSC in the analysis of variables possibly related to change in the achievement tests, but it was impossible to do this because not enough students had complete

data on both the TSC and the achievement tests.

A final recommendation involves the problems encountered in interpreting the results of personality tests administered to persons with a cultural background differing from that of the group on whom the test was standardized. Although indications from the present data are that the AIS students were probably able to respond to the test in a consistent and reliable fashion (i.e., their variability on the subscales was comparable to that of the norm group, and the group profiles were essentially similar on pre- and post-testing), the meaning of the subtest scores remains a question. One hesitates to make the statements that could be made if the present results were obtained from a group of college students, that is, that as a group the AIS students appear to be characterized by a variety of personality disturbances. However, because the TSC appears to reliably measure some kind of variables in this group, hopefully personality characteristics, it would be extremely interesting if it could be determined what exactly it does measure in Indian adolescents. An interdisciplinary approach utilizing psychologists and anthropologists would appear to be a fruitful one in this particular instance.

References

Albuquerque Indian School, FY 1971-1972, Title I Proposal: The motivating school community, Mimeographed paper, undated.

Finney, J. C. Culture change, mental health, and poverty. New York: Simon & Schuster, 1970.

Fitts, W. H. Manual: Tennessee Self-Concept Scale. Counsel Recordings and Tests, Nashville, Tennessee, 1965.

McCandless, B. R. Children: Behavior and development. New York: Holt, Rinehart, & Winston, Inc., 1967.

Moeny, W. Student evaluations: Motivational Environment Program. Mimeographed report, 1971.

Morrison, D. F. Multivariate statistical methods. New York: McGraw-Hill Book Co., 1967.

Reedy, R. L. Summary analysis of the FY 1970 Achievement Testing Program at Albuquerque Indian School. Mimeographed report, undated.

Wolman, B. B. Handbook of clinical psychology. New York: McGraw-Hill Book Co., 1965.

Appendix A

Nature and Meaning of TSC Scores[2]

1. The Self Criticism Score (SC). This scale is composed of 10 items.[3] These are all mildly derogatory statements that most people admit as being true for them. Individuals who deny most of these statements most often are being defensive and making a deliberate effort to present a favorable picture of themselves. High scores generally indicate a normal, healthy openness and capacity for self-criticism. Extremely high scores (above the 99th percentile) indicate that the individual may be lacking in defenses and may in fact be pathologically undefended. Low scores indicate defensiveness, and suggest that the Positive Scores are probably artifically elevated by this defensiveness.

2. The Positive Scores (P). These scores derive directly from the phenomenological classification scheme already mentioned. In the original analysis of the item pool the statements seemed to be conveying three primary messages: (1) This is what I am, (2) This is how I feel about myself, and (3) This is what I do. On the basis of these three types of statements the three horizontal categories were formed. They appear on the Score Sheet as Row 1, Row 2, and Row 3 and are hereafter referred to by those labels. The Row Scores thus compromise three subscores which, when added, constitute the Total Positive or Total P Score. These scores represent an internal frame of reference within which the individual is describing himself.

Further study of the original items indicated that they also varied considerably in terms of a more external frame of reference. Even within the

---

[2]This material is taken from the TSC Manual (Fitts, 1965).

[3]These items have been taken from the L-Scale of the Minnesota Multiphasic Personality Inventory (1951), Copyright 1943, the University of Minnesota. Published by the Psychological Corporation. Reproduced by special arrangements.

same row category the statements might vary widely in content. For example, with Row 1 (the What I am category) the statements refer to what I am physically, morally, socially, etc. Therefore, the pool of items was sorted again according to these new vertical categories, which are the five Column Scores of the Score Sheet. Thus, the whole set of items is divided two ways, vertically into columns (external frame of reference) and horizontally into rows (internal frame of reference) with each item and each cell contributing to two different scores.

a. Total P Score. This is the most important single score on the Counseling Form. It reflects the overall level of self-esteem. Persons with high scores tend to like themselves, feel that they are persons of value and worth, have confidence in themselves, and act accordingly. People with low scores are doubtful about their own worth; see themselves as undesirable; often feel anxious, depressed, and unhappy; and have little faith or confidence in themselves.

If the Self Criticism (SC) Score is low, high P Scores become suspect and are probably the result of defensive distortion. Extremely high scores (generally above the 99th percentile) are deviant and are usually found only in such disturbed people as paranoid schizophrenics who as a group show many extreme scores, both high and low.

On the Counseling Form the Positive Scores are simply designated as P Scores, while on the Score Sheet of the C and R Form they are referred to as P + N Scores in order to clarify the computations involved.

b. Row 1 P Score - Identity. These are the "what I am" items. Here the individual is describing his basic identity - what he is as he sees himself.

c. Row 2 P Score - Self Satisfaction. This score comes from those items where the individual describes how he feels about the self he perceives. In

general this score reflects the level of self satisfaction or self acceptance. An individual may have very high scores on Row 1 and Row 3 yet still score low on Row 2 because of very high standards and expectations for himself. Or vice versa, he may have a low opinion of himself as indicated by the Row 1 and Row 3 Scores yet still have a high Self Satisfaction Score on Row 2. The subscores are therefore best interpreted in comparison with each other and with the Total P Score.

d. Row 3 P Scores - Behavior. This score comes from those items that say "this is what I do, or this is the way I act." Thus, this score measures the individual's perception of his own behavior or the way he functions.

e. Column A - Physical Self. Here the individual is presenting his view of his body, his state of health, his physical appearance, skills, and sexuality.

f. Column B - Moral-Ethical Self. This score describes the self from a moral-ethical frame of reference--moral worth, relationship to God, feelings of being a "good" or bad" person, and satisfaction with one's religion or lack of it.

g. Column C - Personal Self. This score reflects the individual's sense of personal worth, his feelings of adequacy as a person, and his evaluation of his personality apart from his body or his relationships to others.

h. Column D - Family Self. This score reflects one's feelings of adequacy, worth, and value as a family member. It refers to the individual's perception of self in reference to his closest and most immediate circle of associates.

i. Column E - Social Self. This is another "self as perceived in relation to others" category but pertains to "others" in a more general way. It reflects the person's sense of adequacy and worth in his social interaction with other people in general.

3. The Variability Scores (V). The V scores provide a simple measure of the

amount of variability, or inconsistency, from one area of self perception to another. High scores mean that the subject is quite variable in this respect while low scores indicate low variability which may even approach rigidity if extremely low (below the first percentile).

a. Total V. This represents the total amount of variability for the entire record. High scores mean that the person's self concept is so variable from one area to another as to reflect little unity or integration. High scoring persons tend to compartmentalize certain areas of self and view these areas quite apart from the remainder of self. Well integrated people generally score below the mean on these scores but above the first percentile.

b. Column Total V. This score measures and summarizes the variations within the columns.

c. Row Total V. This score is the sum of the variations across the rows.

4. The Distribution Score (D). This score is a summary score of the way one distributes his answers across the five available choices in responding to the items of the Scale. It is also interp-eted as a measure of still another aspect of self perception: certainty about the way one sees himself. High scores indicate that the subject is very definite and certain in what he says about himself while low scores mean just the opposite. Low scores are found also at times with people who are being defensive and guarded. They hedge and avoid really committing themselves by employing "3" responses on the Answer Sheet.

Extreme scores on this variable are undesirable in either direction and are most often obtained from disturbed people. For example, schizophrenic patients often use "5" and "1" answers almost exclusively, thus creating very high D Scores. Other disturbed patients are extremely uncertain and noncommittal in their self descriptions with a predominance of "2", "3" and "4" responses and very low D Scores.

5. The Time Score. This score is simply a measure of the time, to the nearest minute, that the subject requires to complete the Scale. The author has only recently made any study of this variable, and at this point little is known as to its meaning or significance. It correlates significantly with only one of the many other scores of the Scale (Net Conflict sub-score for Column C where r=.32, significant at the .05 level). Therefore, any validity it may prove to have with other criteria should add to the total validity of the Scale.

The data do indicate that, provided the individual has sufficient education, intelligence, and reading ability to handle this task, the majority of subjects complete the Scale in less than 20 minutes. These qualifications are quite important; if they are not met, the Time Score obviously has little meaning. It has been found that psychiatric patients in general take longer than nonpatients. This is particularly true of those who are overly compulsive, paranoid, or depressed.

6. The True-False Ratio (T/F). This is a measure of response set or response bias, an indication of whether the subject's approach to the task involves any strong tendency to agree or disagree regardless of item content (Fitts, 1961).

The actual meaning of T/F can be approached in three ways. (1) It can be considered solely as a measure of response set and interpreted in terms of the findings about the meaning of deviant response sets. (2) It can be treated purely as a task approach or behavioral measure which has meaning only in terms of empirical validity. In this sense the T/F Ratio differentiates patients from nonpatients and correlates significantly with other tests. (3) It can also be considered from the framework of self theory. From this approach, high T/F Scores indicate the individual is achieving self definition or self description by focusing on what he is and is relatively unable to accomplish the same thing by eliminating or rejecting what he is not. Low T/F Scores would mean the exact opposite, and scores in this middle

range would indicate that the subject achieves self definition by a more balanced employment of both tendencies--affirming what is self and eliminating what is not self.

7. Net Conflict Scores. These scores are highly correlated with the T/F Score. More directly, however, they measure the extent to which an individual's responses to positive items differ from, or conflict with, his responses to negative items in the same area of self perception. Thus, this is a limited and purely operational definition and application of the term "conflict." On the C and R Score Sheet separate scores are computed within each cell for the positive and negative items. The difference between these scores, the P - N Score, is an operational measure of conflict. Since the responses on the negative items are reversed on the Score Sheet, the P Scores and the N Scores have equivalent meanings. Thus, any difference between P and N reflects contradiction or conflict.

There are two different kinds of conflict, as follows:

a. Acquiescense Conflict. This phenomenon occurs when the P Scores are greater than the N Scores (P - N yields a positive score or number). This means that the subject is overaffirming his positive attributes.

b. Denial Conflict. This is the opposite of acquiescense conflict. Here the N Scores for the cells are higher than the P Scores (P - N yields minus scores). This means that the subject is overdenying his negative attributes in relation to the way he affirms his positive characteristics. He concentrates on "eliminating the negative."

8. Total Conflict Scores. The foregoing Net Conflict Scores were concerned only with directional trends in our P - N measure of conflict. However, some individuals have high P - N differences which cancel each other out because they are so variable in direction. It is of equal interest to determine the

total amount of P - N conflict in a subject's self concept as well as the net or directional amount of conflict. The total Conflict score does this by summing P - N discrepancies regardless of sign. High scores indicate confusion, contradiction, and general conflict in self perception. Low scores have the opposite interpretation, but extremely low scores (below the red line on the Profile Sheet) have a different meaning. The person with such low scores is presenting such an extremely tight and rigid self description that it becomes suspect as an artificial, defensive stereotype rather than his true self image. Disturbed people generally score high on this variable, but some also have deviantly low scores depending on the nature and degree of their disorder.

The conflict scores are reflections of conflicting responses to positive and negative items within the same area of self perception. These scores are not to be confused with the variability scores, which reflect fluctuations from one area of self perception to another.

9. The Empirical Scales. These six scales were all derived by item analysis, with a resulting selection of those items which differentiated one group of subjects from all other groups. The scores on these scales are purely empirical, and cut across the basic classification scheme of the Scale.

These scales were derived from an analysis of item responses with the following groups:

Group	Size of Group
Norm Group	626
Psychotic Group (Psy)	100
Neurotic Group (N)	100
Personality Disorder Group PD)	100
Defensive Positive Group (DP)	100
Personality Integration Group (PI)	75

The comparative item responses for these groups were studied and analyzed by Chi Square tests. Those items which differentiated one group from all other

groups were then used to compose a specific scale for that group. There is some overlapping of items, since a number of items are used on more than one scale.

The six empirical scales derived by this method, in order of their appearance on the Profile Sheet, are as follows:

a. The Defensive Positive Scale (DP). This is a more subtle measure of defensiveness than the SC Score. One might think of SC as an obvious defensiveness score and DP as a subtle defensiveness score. The DP Score stems from a basic hypothesis of self theory: that individuals with established psychiatric difficulties do have negative self concepts at some level of awareness, regardless of how positively they describe themselves on an instrument of this type.

With this basic assumption, the author collected data on 100 psychiatric patients whose Total P Scores were above the mean for the Norm Group. The item analysis then identified 29 items which differentiated this DP Group from the other groups.

The DP Score has significance at both extremes. A high DP Score indicates a positive self description stemming from defensive distortion. A significantly low DP Score means that the person is lacking in the usual defenses for maintaining even minimal self esteem.

b. The General Maladjustment Scale (GM). This scale is composed of 24 items which differentiate psychiatric patients from nonpatients but do not differentiate one patient group from another. Thus, it serves as a general index of adjustment-maladjustment but provides no clues as to the nature of the pathology. Note that this is an inverse Scale on the Profile Sheet. Low raw scores result in high T-Scores, and vice versa.

c. The Psychosis Scale (Psy). The Psy Scale is based on 23 items which

best differentiate psychotic patients from other groups.

d. The Personality Disorder Scale (PD). The 27 items of this scale are those that differentiate this broad diagnostic category from the other groups. This category pertains to people with basic personality defects and weaknesses in contrast to psychotic states or the various neurotic reactions. The PD Scale is again an inverse one.

e. The Neurosis Scale (N). This is an inverse scale composed of 27 items. As with other inverse scales, high T-Scores on the Profile Sheet still mean high similarity to the group from which the scale was derived--in this case neurotic patients.

f. The Personality Integration Scale (PI). The scale consists of the 25 items that differentiate the PI Group from other groups. The scoring is slightly different for this scale and is explained on the special template for scoring this scale. This group was composed of 75 people who, by a variety of criteria, were judged as average or better in terms of level of adjustment or degree of personality integration.

10. The Number of Deviant Signs Score (NDS). The NDS Score is a purely empirical measure, and is simply a count of the number of deviant features on all other scores. This score is based upon the theoretical position of Berg (1957) as stated in his "deviation hypothesis." This hypothesis states that individuals who deviate sharply from the norm in minor behaviors are likely to be deviant in more major aspects of behavior. The findings with the NDS Score substantiate this hypothesis. Disturbed persons often obtain extreme scores on either end of the continuum. Consequently, a system which sets appropriate cut-off points for each score on the Scale will identify disturbed persons with considerable accuracy.

The NDS Score is the Scale's best index of psychological disturbance. This score alone identifies deviant individuals with about 80% accuracy.

Section III

Analysis of Behavioral Data

Daniel B. Berch, Ph.D.

Analysis of Behavioral Data

The basic purpose of this section of the report is to examine the behavioral data in an effort to evaluate the effects of the MEP on the particular target behaviors which it was designed to improve.

Design of the study. As noted in Blanchard's statement at the beginning of this report, the MEP did not involve a formal experimental design. Although no control group was used, the MEP inadvertently implemented what is known as the "reversal technique." This is a type of within-subjects design in which each student serves as his own control. The basic design is as follows: (1) Baseline Period--in which the naturally occurring behavior is measured; (2) Experimental Period$_1$--introduction of treatment variables; (3) Reversal Period--return to baseline conditions; (4) Experimental Period$_2$--reinstatement of treatment variables.

For the MEP, behavioral measures were taken from 8/24/70 to 10/11/70, prior to the introduction of the token economy. This time interval would correspond to the Baseline Period. The first phase of the token economy was implemented on 10/12/70 and continued to operate through December 1970. This period corresponds to the Experimental Period$_1$ and will be referred to as Phase 1 of the token program throughout the rest of this section. It is the author's understanding that the MEP was not reintroduced in January 1971 in order to allow time for a reduction of tensions between the staff and the administration, since the staff had demonstrated considerable hostility toward the token economy. The token program was not reinstated until the last week of March 1971. Thus, the period from January 1971 to the end of the third week of March 1971 may be considered a Reversal Period.

However, even though some of the teachers and dormitory aides began dispensing tokens during the last week of March 1971, others did not begin until the first week of April 1971. Therefore, the data from the last week of March 1971 were excluded from the analyses.

The token program was in operation from April 1971 to the end of the third week in May 1971. This period corresponds to the second experimental period, i.e., Phase 2 of the token program. In order to determine whether the token program was having any effect on the target behaviors, one must determine whether changes in the rate of behavior correspond to changes in the conditions of reinforcement. Where records were available for the Baseline and Phase 1 periods, individual curves were plotted in order to determine whether the reinforcers were influencing the target behaviors.

Since most of the data made available to the evaluation team covered the period from January 1971 to May 1971, it became apparent that the reversal design could not be considered the most accurate description of the MEP study. Actually, the design of this study is characterized better by what Campbell and Stanley (1963) refer to as a "Time-Series Design." This is a type of "quasi-experimental" design in the sense that although aspects of a true experiment were utilized in the data collection procedures, full control over the dispensing of tokens at the appropriate time and place was lacking. From the standpoint of a Time-Series Design, to determine whether the MEP did in fact influence behavioral change, one must examine the outcome patterns generated by the results of the series of measurements. If there is a discontinuity in the trend following implementation of the token economy, one would have evidence that the MEP was effective.

Sources of data. The data were gathered primarily from the dormitories and the classrooms. As it turned out, reasonably complete records were avail-

able for only five behaviors: Drinking, AWOL, Tardy (bedcheck), Absence (class), and Tardy (class). Most of these records covered the period from January 1971 through May 1971. A midterm evaluation of the MEP is available which assesses the data from late August 1970 through December 1970; unfortunately, the raw data from this period was not retained for the present evaluation team. Moreover, most of the original records were discarded by the staff at the Indian School. As it turned out, the only relatively complete data available for the period from August 1970 to December 1970 consisted of records on Drinking for Tanoan Hall and Drinking and AWOL for Wauneka Hall.

All the data were analyzed with respect to 28-day periods, corresponding roughly to the months from January 1971 to May 1971. The rationale for this procedure arose from the notion of the Time-Series Design discussed above. That is, in order to be able to make conclusions regarding causal relationships between the MEP and behavioral change, enough data points must exist to determine any trends. However, since most of the students did not emit the target behaviors daily nor even weekly during the Reversal Period, the decision was made that it would be meaningless to look at daily or weekly rates. Anything larger than monthly rates, however, would not provide enough data points to determine any trends.

For the months in which some daily records were missing, the data were transformed so as to represent a 28-day rate. For example, Wauneka Hall had records for only two weeks of January 1971 for most of the dormitory behaviors. Thus, if a girl was caught drinking three times during that period, her score was converted to "six" for the month of January.

Approaches to assessment of behavioral change. There were three main approaches to the analysis of the behavioral data. One of these involved

the assessment of the percentage of students emitting a particular target behavior at least once during each month, from January 1971 to May 1971. The second manner of treating the data was to determine for each dormitory the frequency (total number of instances) of the target behaviors for each month from January through May. The third approach involved the plotting of curves for certain individual students, which illustrated the number of instances of the target behaviors occurring each month. This last approach included records from August 1970 through December 1970 where available.

Percentage data. Figures 1 through 9 represent graphically the percentage of students who emitted the designated target behaviors at least once during each month from January 1971 through May 1971. The data are categorized by dormitory, AIS or Bordertown,[1] as well as by sex for Kiva Hall. The McNemar test for significance of change was performed on the percentage data. This test was used to determine the reliability of apparent increases and decreases between two of the monthly periods in the percentage of students emitting each particular behavior. In most cases, comparisons were made between January and May. When the percentage was lower in February than in January, February and May were compared. When the trend appeared to change following reinstatement of the token program, the percentages from January and April were compared, followed by another comparison between April and May. A summary of the significant changes is presented in Table 1. One may note that in some cases, relatively large percentage changes were less highly significant than much lower percentage changes. Closer examination of Table 1 will reveal that these high percentages represent changes for Kiva Hall, which has a much smaller number of students than either Tanoan or Wauneka. With the McNemar

[1]Bordertown refers to students who attend schools off campus.

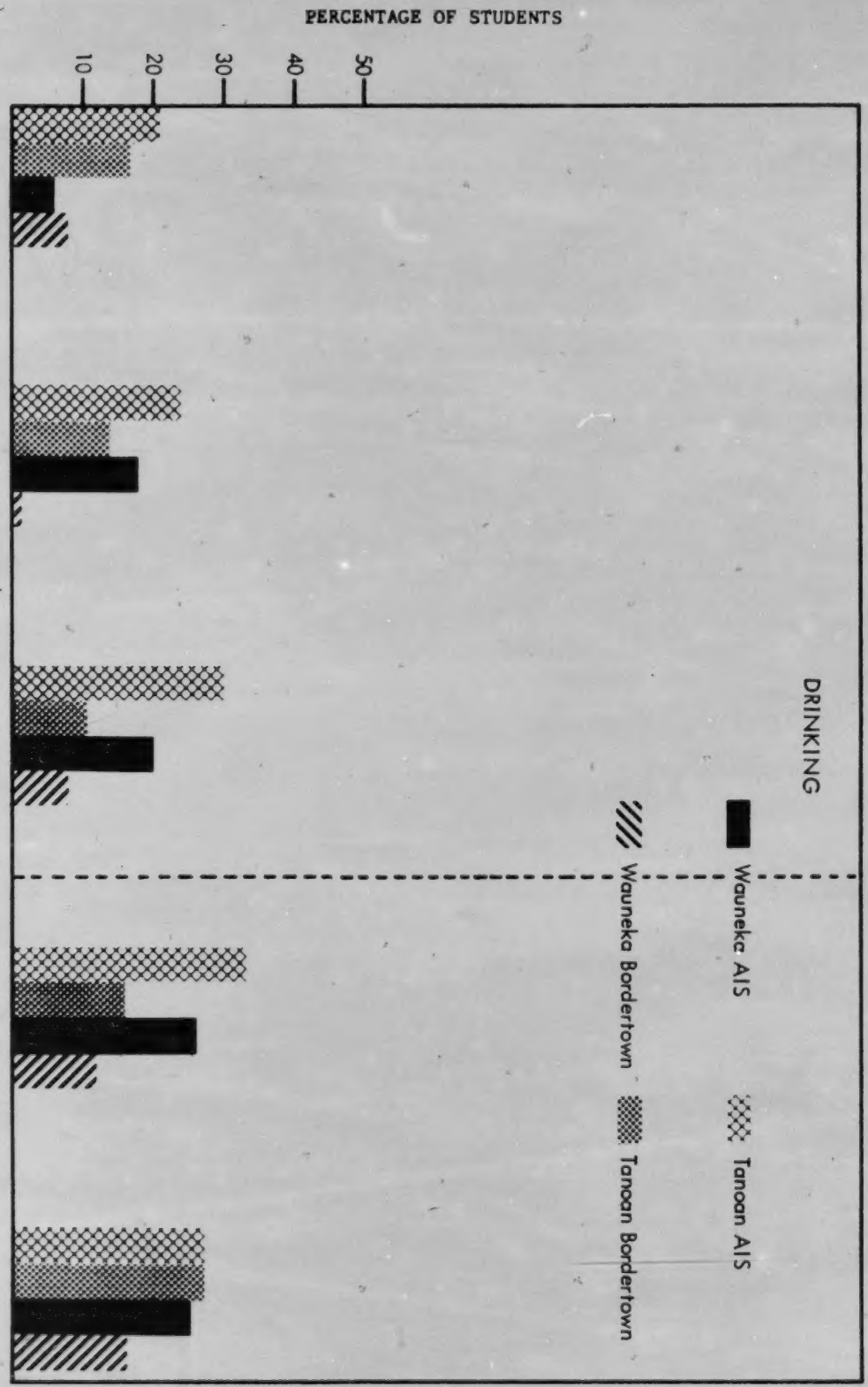

Fig. 1 Percentage of students drinking at least once each month.

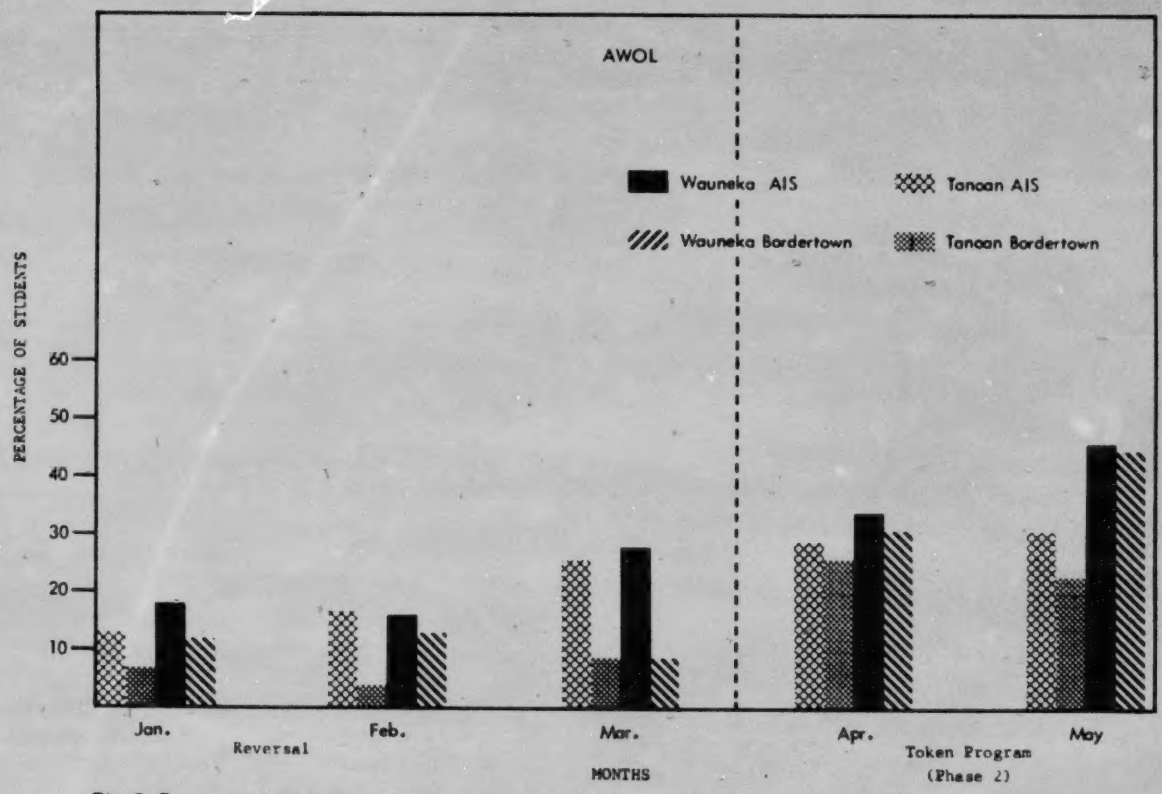

Fig. 2 Percentage of students **AWOL** at least once each month.

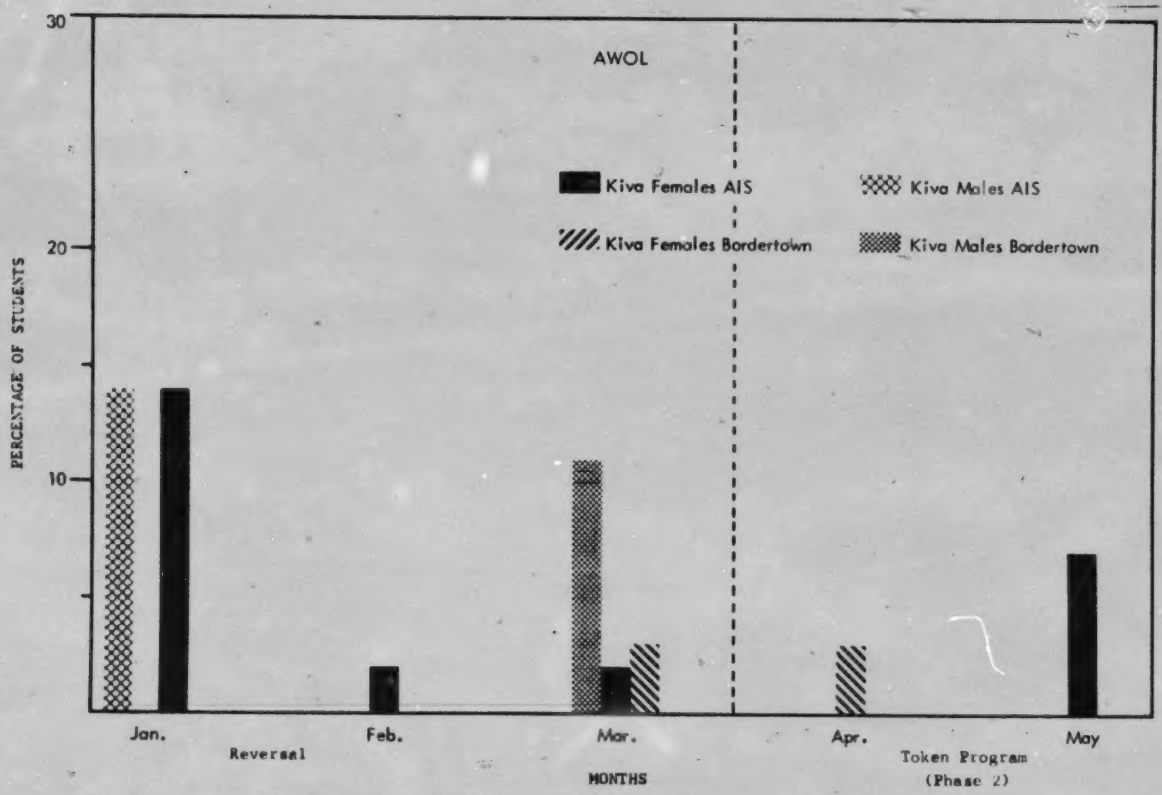

Fig. 3 Percentage of students **AWOL** at least once each month.

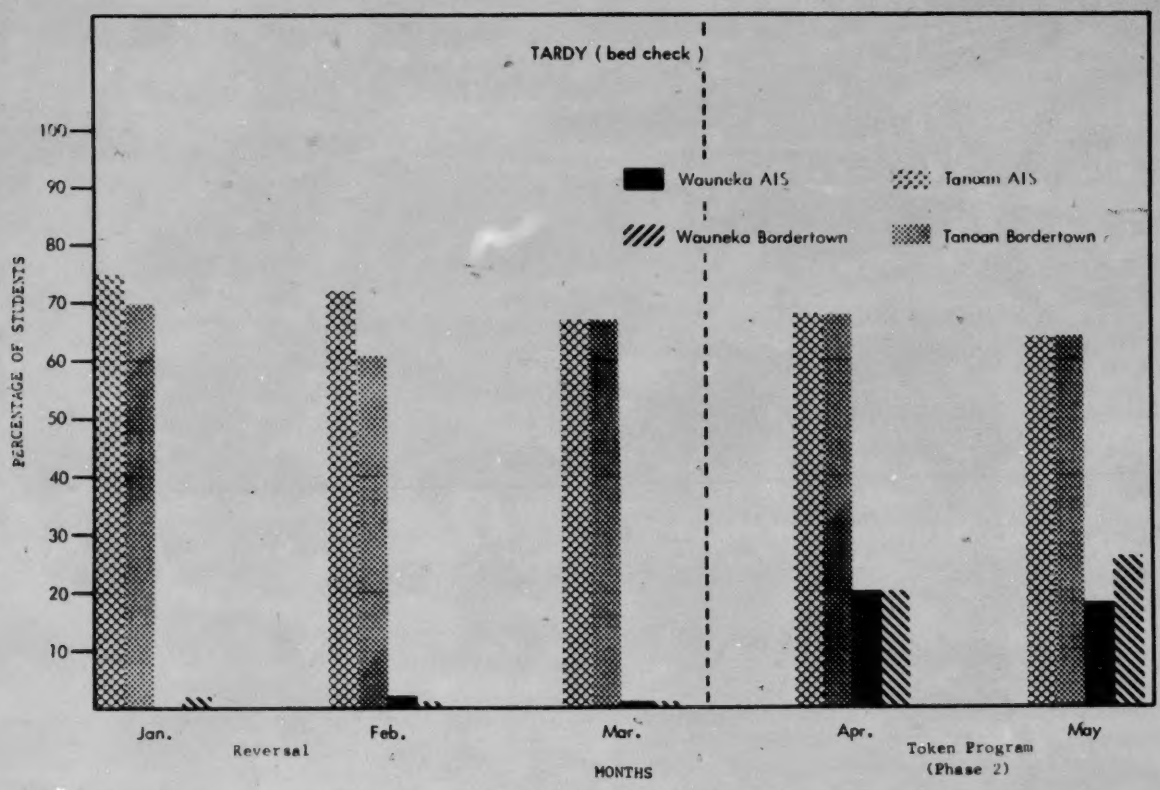

Fig. 4 Percentage of students **tardy (bedcheck)** at least once each month.

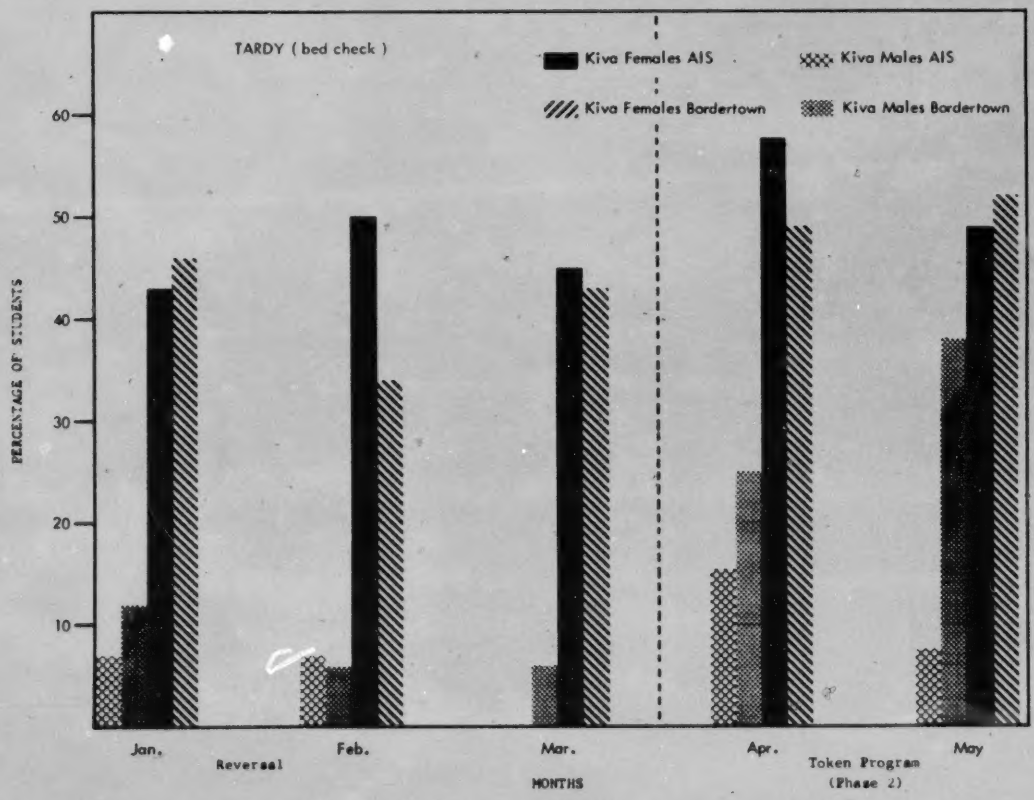

Fig. 5 Percentage of students **tardy (bedcheck)** at least once each month.

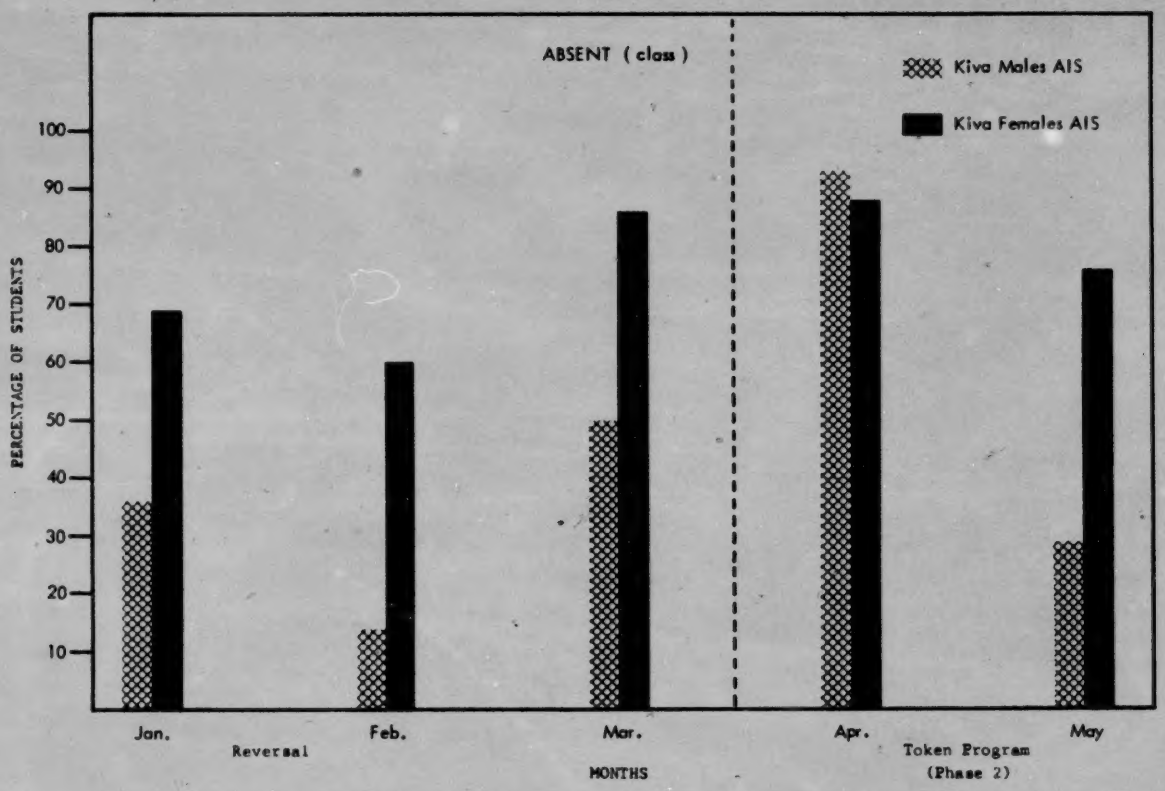

Fig. 6 Percentage of students absent (class) at least once each month.

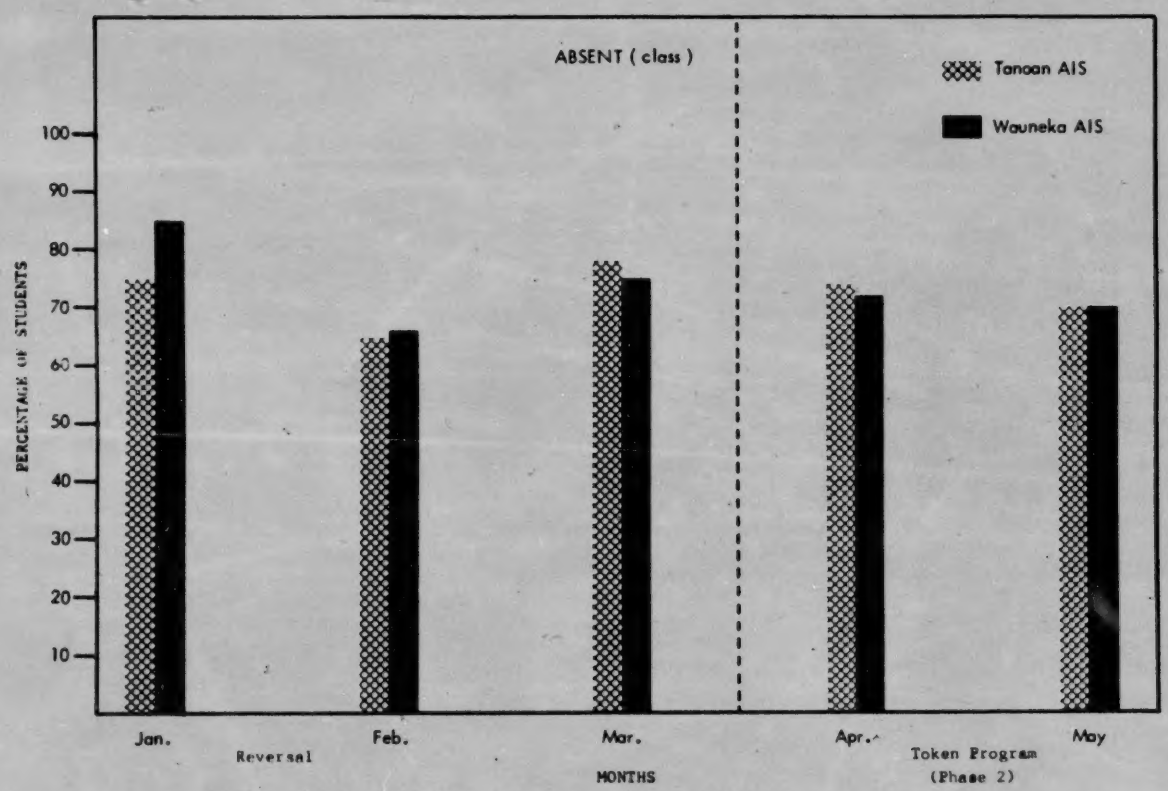

Fig. 7 Percentage of students absent (class) at least once each month.

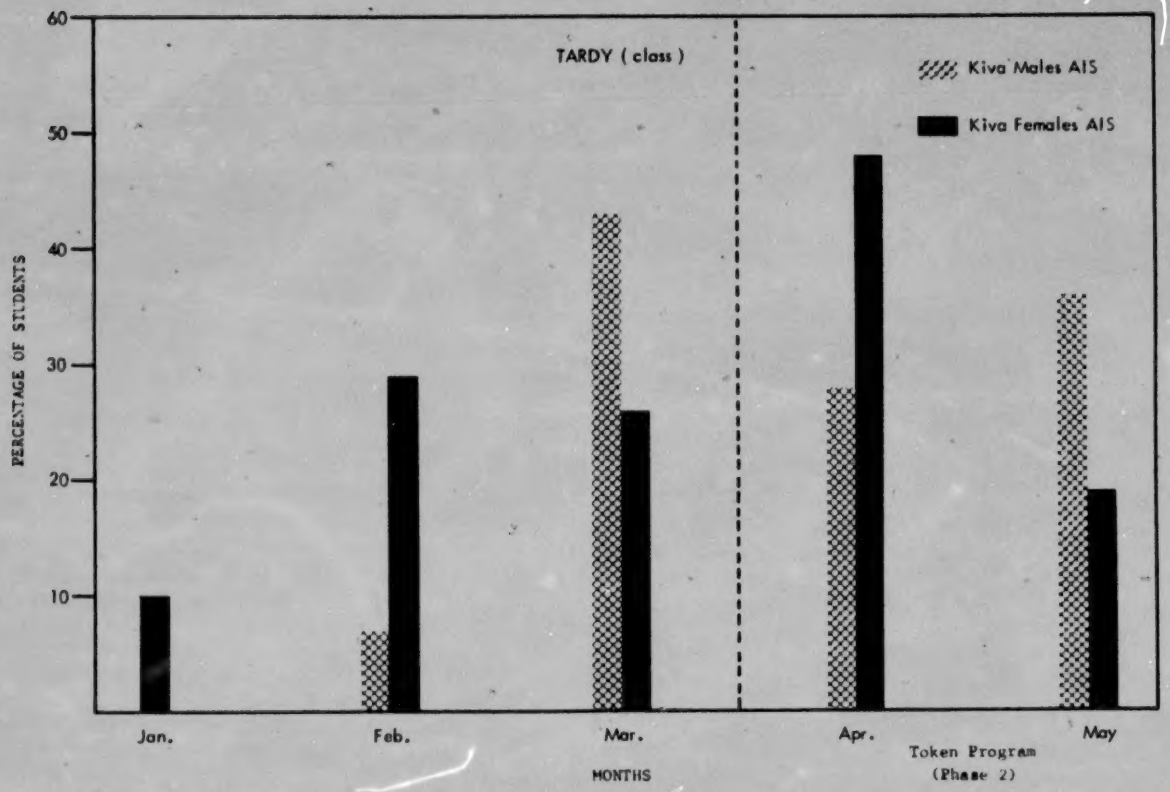

Fig. 8 Percentage of students tardy (class) at least once each month.

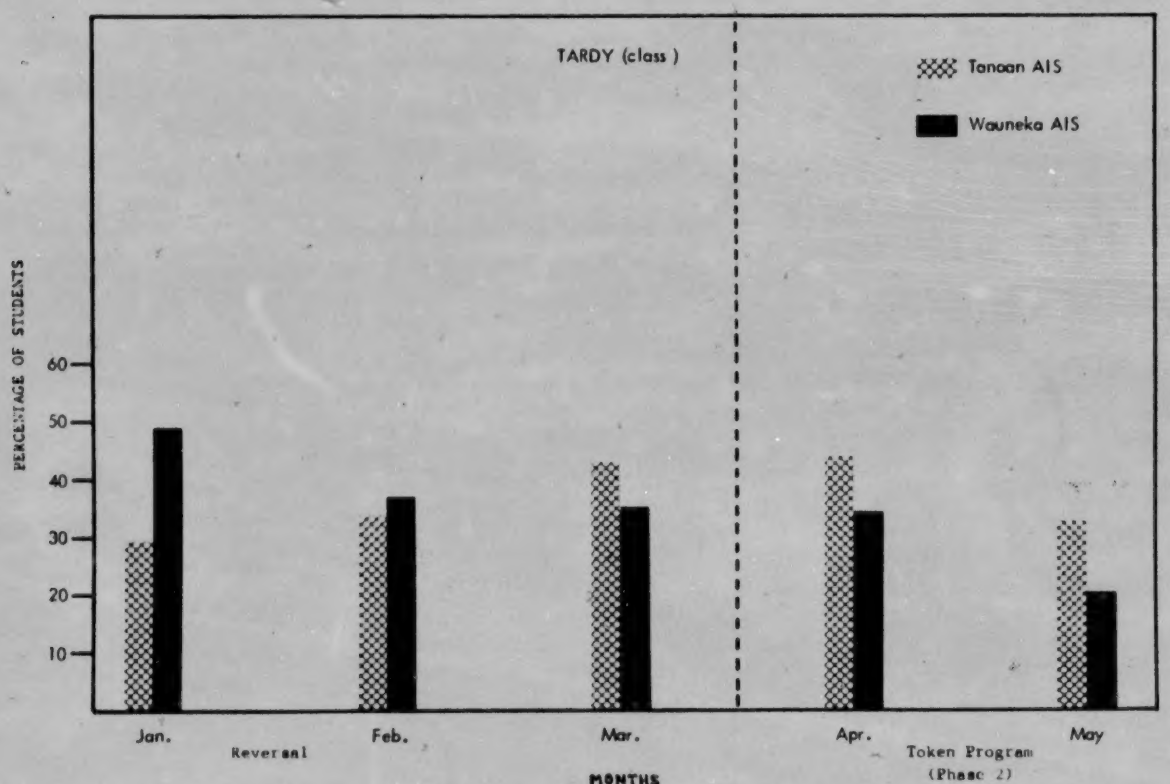

Fig. 9 Percentage of students tardy (class) at least once each month.

test, when the number of persons emitting the behavior is rather small, relatively large percentage changes may still not be highly significant. As shown in Table 1, most of the significant changes represent increases in the percentage of students emitting the target behaviors. The only significant decreases in percentage were for the following: Tardy (bedcheck) for Tanoan AIS; Absence (class) for Wauneka AIS and Kiva Males AIS; Tardy (class) for Tanoan AIS, Wauneka AIS and Kiva Females AIS. Of these six changes, only three represent an unequivocal discontinuity in trend after reinstatement of the token program: 1) Absence (class) for Kiva Males AIS, 2) Tardy (class) for Tanoan AIS, and 3) Tardy (class) for Kiva Females AIS.

Frequency data. The data were also analyzed in terms of the total number of instances of the target behaviors occurring each month from January through May. In most cases the total number of instances for each month is plotted. The means were not used, except for Absence (class) data, because they were less than one and thus would be rather meaningless to the reader. For all five behaviors, statistical analyses were run to determine: 1) whether the patterns of the dormitories differed significantly from each other, i.e., that they were not parallel, and 2) whether particular trends departed from a horizontal line, i.e., whether there were reliable increases or decreases.

Figure 10 shows the frequency of Drinking for Tanoan Hall and Wauneka Hall from January through May. For Tanoan Hall, Drinking increased from January through March, decreased somewhat in April, and increased again in May. The Drinking pattern for Wauneka consisted of an increasing trend from January through May. A statistical analysis of trend indicated that the general increasing patterns for Tanoan and Wauneka are reliable ($F = 4.31$; $df = 4, 613$; $p < .005$). Furthermore, the patterns of the two dormitories do not differ significantly ($F = 1.95$, $df = 4, 613$; $p > .05$).

Table 1

Summary of the Significant Changes in the Percentage of Students Emitting a Designated Behavior at Least Once During a Particular Month

Behavior	Dorm	Months	Percentage Change	Direction of change	Significance Level
Drinking	Wauneka AIS	Jan & May	19	(+)Increase	.0003
AWOL	Tanoan AIS	Jan & May	18	(+)Increase	.00005
AWOL	Tanoan Bdtn.	Jan & May	16	(+)Increase	.0004
AWOL	Wauneka AIS	Jan & May	28	(+)Increase	.00005
AWOL	Wauneka Bdtn.	Jan & May	33	(+)Increase	.00003
Tardy (bedcheck)	Tanoan AIS	Jan & May	11	(-)Decrease	.03
Tardy (bedcheck)	Wauneka AIS	Feb & May	16	(+)Increase	.0005
Tardy (bedcheck)	Wauneka Bdtn.	Jan & May	24	(+)Increase	.00003
Absent (class)	Wauneka AIS	Jan & May	16	(-)Decrease	.003
Absent (class)	Kiva Males AIS	Jan & April	57	(+)Increase	.03
Absent (class)	Kiva Males AIS	April & May	64	(-)Decrease	.008
Absent (class)	Kiva Females AIS	Jan & April	19	(+)Increase	.02
Tardy (class)	Tanoan AIS	Jan & April	16	(+)Increase	.00005
Tardy (class)	Tanoan AIS	April & May	11	(-)Decrease	.0003
Tardy (class)	Wauneka AIS	Jan & May	31	(-)Decrease	.00006
Tardy (class)	Kiva Females AIS	Jan & April	38	(+)Increase	.0005
Tardy (class)	Kiva Females AIS	April & May	29	(-)Decrease	.003

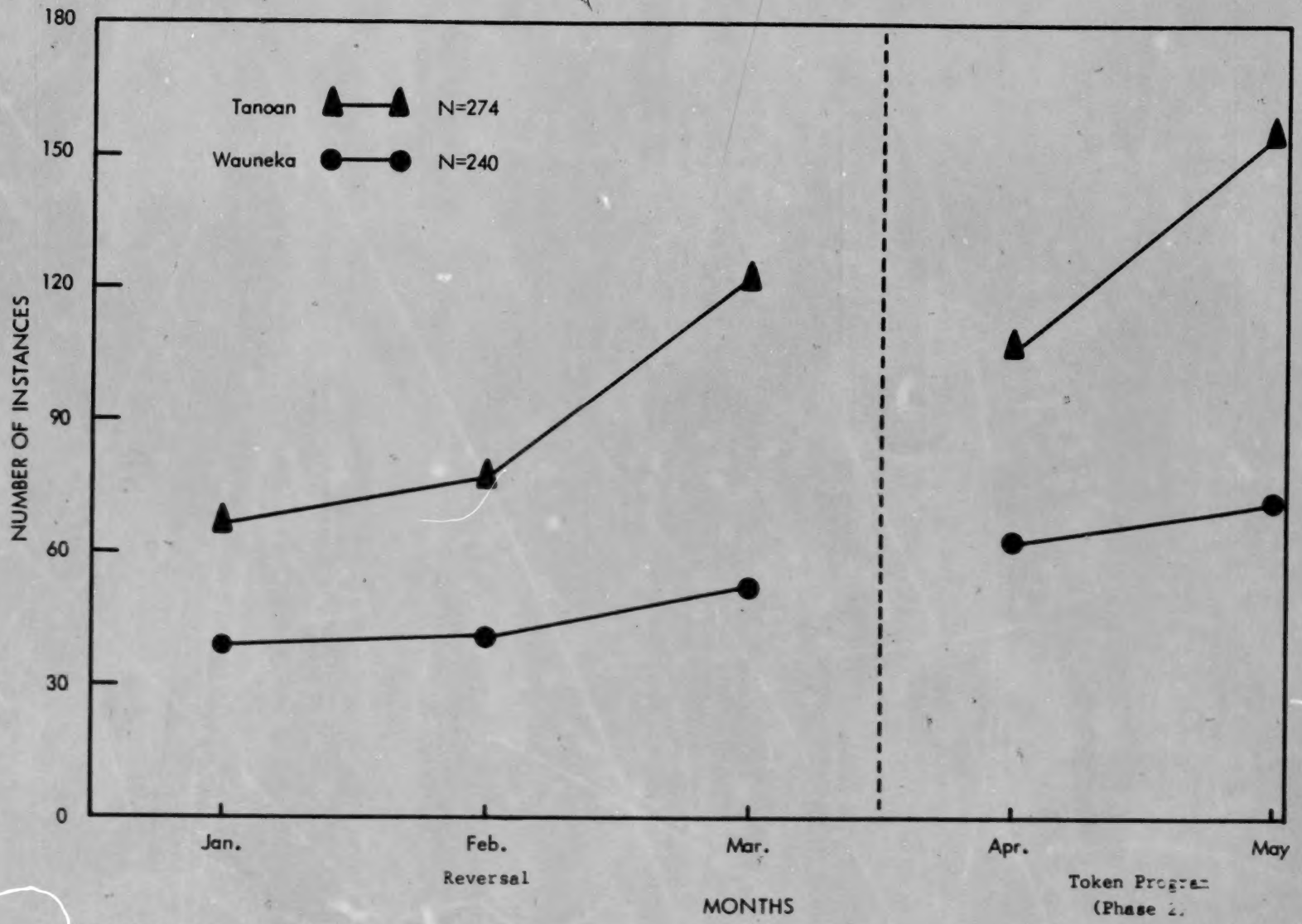

Fig.10 Total number of instances of drinking each month for TANOAN and WAUNEKA.

With respect to AWOL, Figure 11 shows that the number of instances of this behavior increased for Tanoan Hall during the Reversal Period, but decreased somewhat during Phase 2 of the token program. Although the number of AWOLs increased moderately for Wauneka Hall during the Reversal Period, there was a marked increase during April and May. There appears to be negligible change for Kiva Hall from January through May. Analysis of trend for these data indicated that the trends depart reliably from a horizontal line ($f = 2.62$; $df = 4, 613$; $p < .05$). The patterns of the dormitories differ significantly as well ($F = 1.96$; $df = 8, 1226$; $p < .05$), which suggests that there was essentially no change in trend for Kiva Hall.

The data on Tardy (bedcheck) for Tanoan Hall and Kiva Hall are presented in Figure 12. The frequency of Tardy (bedcheck) increased for Tanoan Hall through March, decreased in April, and increased markedly in May. There was a general increasing trend for Kiva Hall. The data from Wauneka is not included because it was discovered that the available records grossly underestimated the frequency of this behavior. Since this error was discovered after the analyses had been run, the results of the statistical analyses cannot be considered accurate. Nevertheless, the marked similarity between the patterns for Tardy (bedcheck) and the patterns for Drinking suggest that as in the case of Drinking, the patterns of Tardy (bedcheck) are not significantly different for Tanoan and Kiva, but both trends differ reliably from a horizontal line.

Figure 13 presents the mean number of instances of Absence (class) for the three dorms. All three showed an increasing trend from February through March. There was a decrease for Tanoan and Wauneka from March to April, and a slight increase for Kiva. However, all three dormitories showed decreases from April to May. Both the departure from flatness and the differences among the patterns of the dormitories were highly significant ($F = 38.98$; $df = 4, 328$;

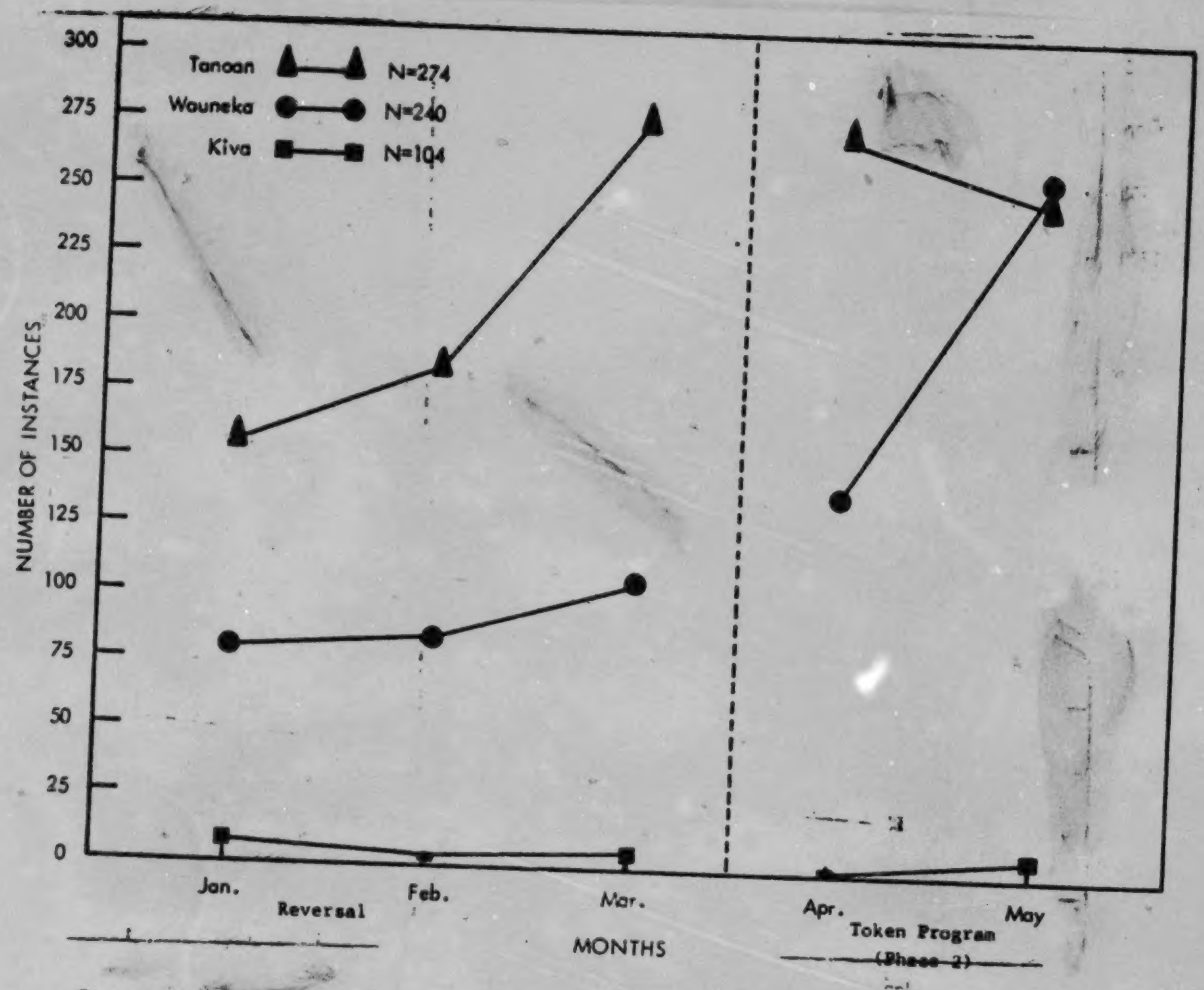

Fig.11 Total number of instances of AWOL each month for each of the three dormitories.

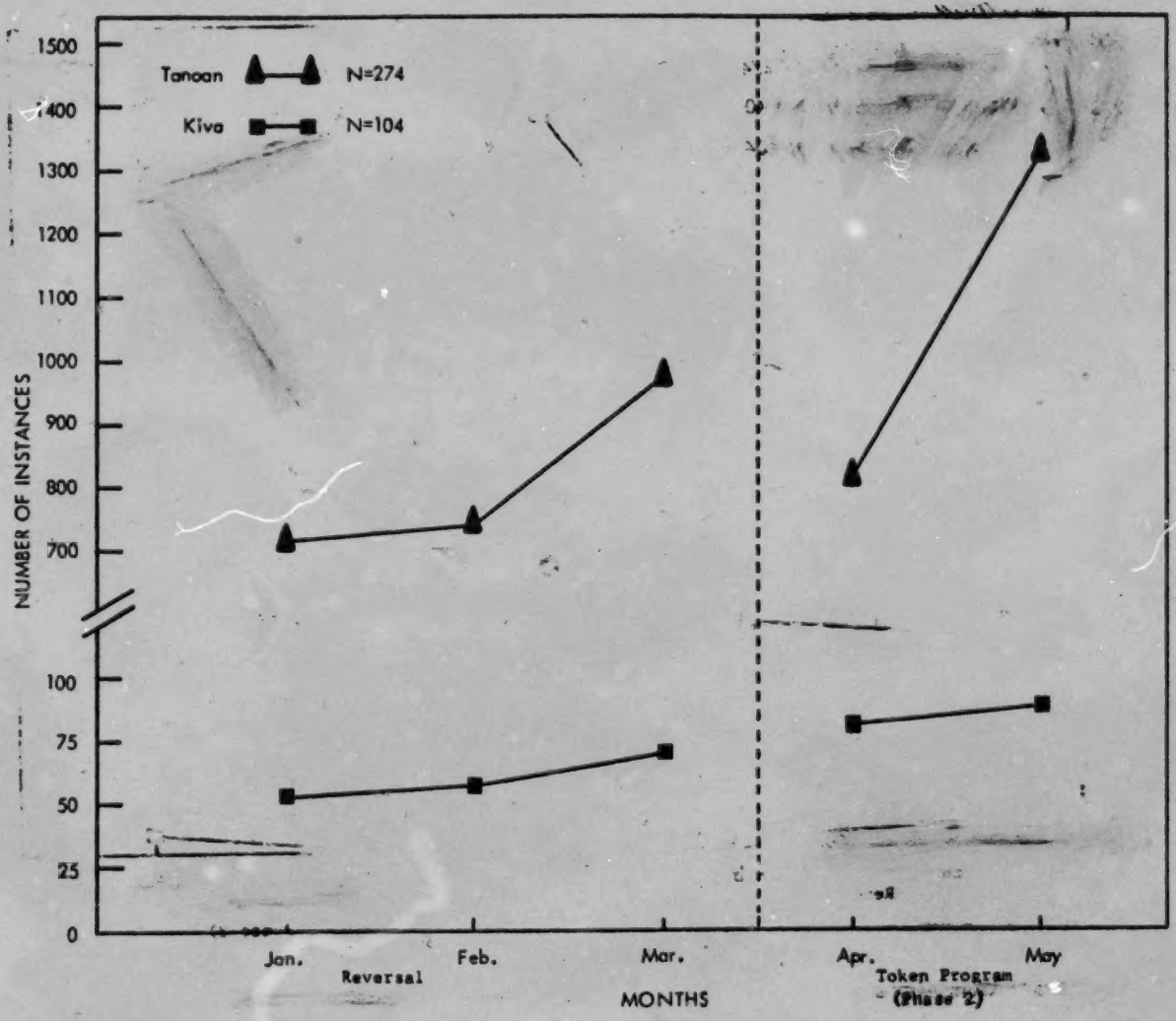

Fig.12 Total number of instances of tardiness (bedcheck) each month for TANOAN and KIVA.

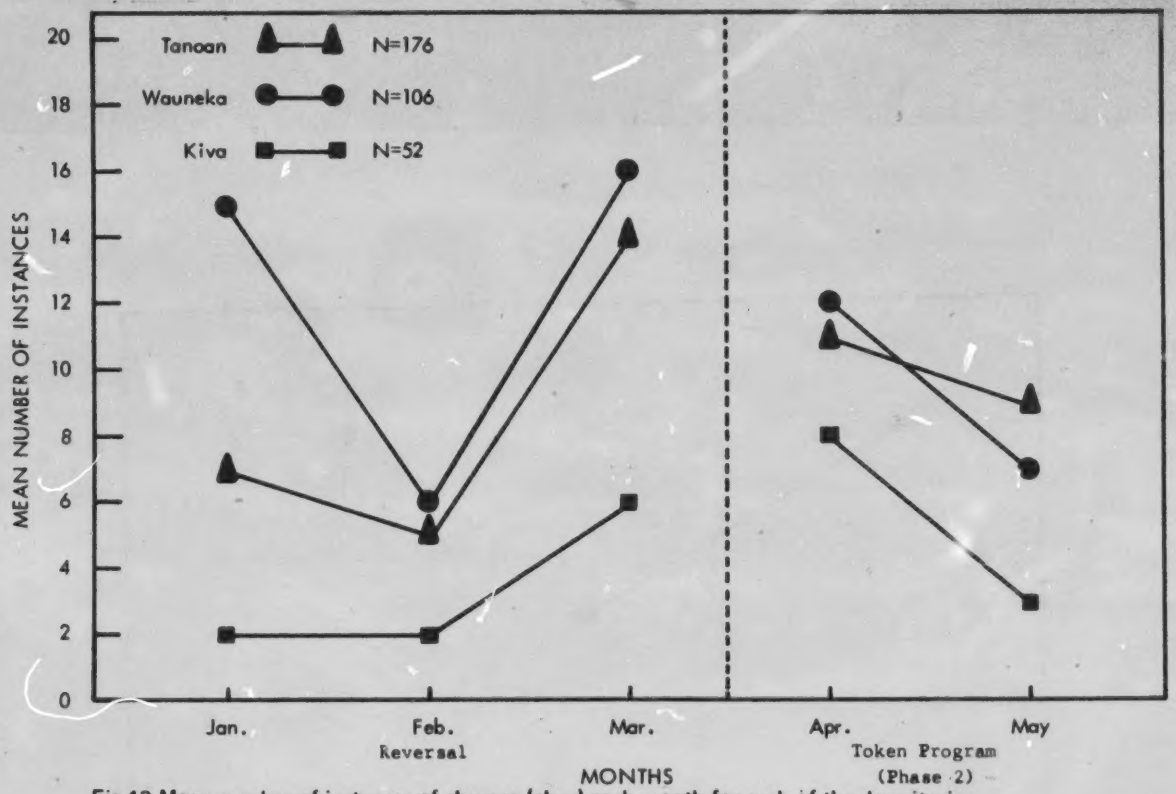

Fig.13 Mean number of instances of **absence (class)** each month for each of the dormitories.

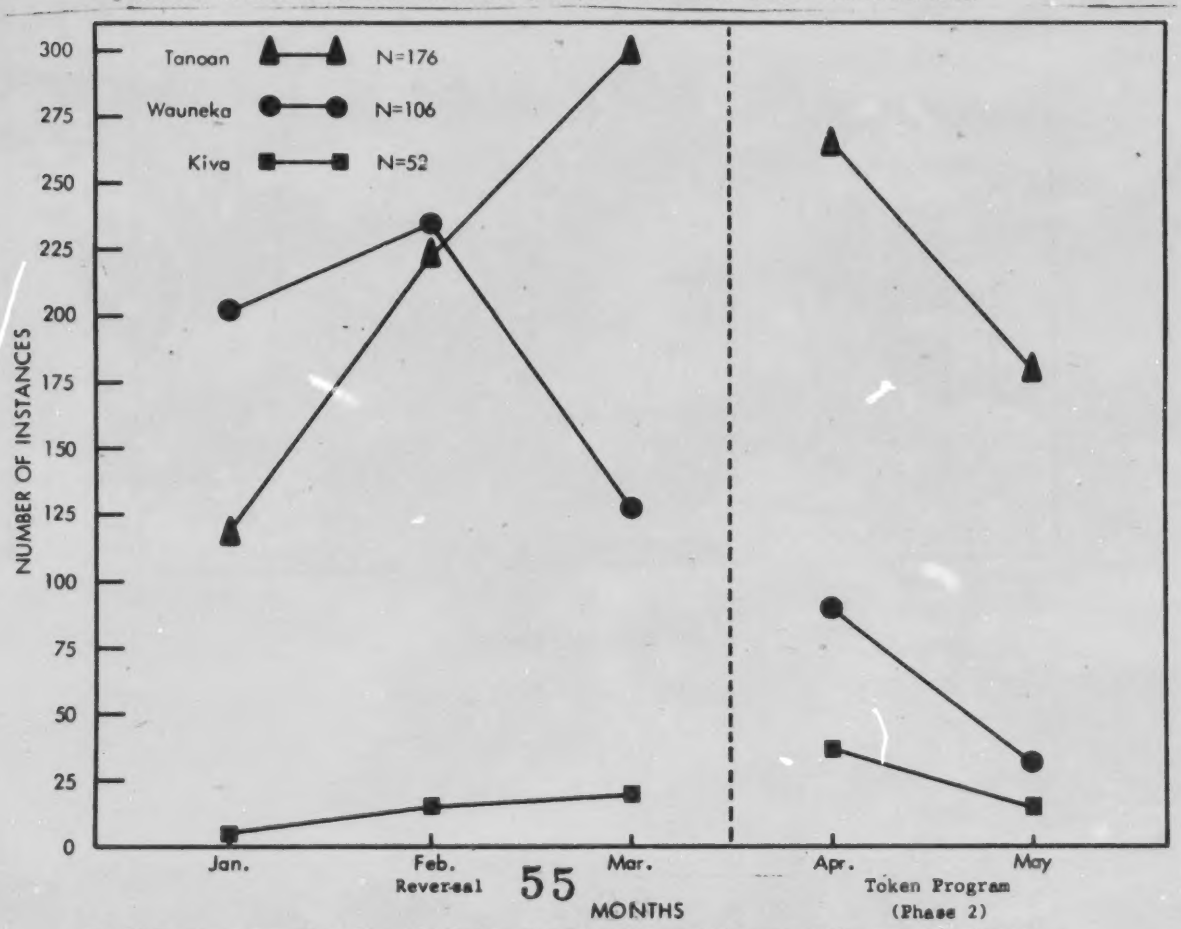

Fig.14 Total number of instances of **tardiness (class)** each month for each of the three dormitories.

$p < .001$ and $F = 7.19$; $df = 3, 656$; $p < .001$, respectively).

Figure 14 illustrates the rate of Tardy (class) for all three dorms. Tanoan Hall showed an increase during the Reversal Period, then decreased in April and May. Wauneka decreased from February to March and continued to decrease through April and May. Kiva increased through April and decreased in May. The patterns of the dormitories were found to differ significantly ($F = 7.50$; $df = 8, 656$; $p < .001$). In addition, the trends depart reliably from a horizontal line ($F = 4.82$; $df = 4, 328$; $p < .001$).

Individual curves. Since the percentage data indicated that a number of students did not emit certain behaviors, it is likely that the averaging process distorted changes in behavior that took place for individuals. In keeping with a behavior modification approach to evaluation, one should look at individual curves in order to determine whether the reinforcement contingencies were influencing behavioral change. One difficulty, however, involves the basis for selection of individual curves. It was decided that some of the data from the student questionnaire should be used as a basis for selecting individuals. Question 41 of the student questionnaire (see Landau) investigated the students' perceptions of changes in their behavior following implementation of the MEP. Most students perceived little change in their behaviors. However, a few did perceive extreme changes. It follows that these same changes should be evident in the behavioral data. Consequently, individual curves were plotted for the behaviors which were perceived as having undergone extreme change following implementation of the MEP.

As shown in Figure 15, Student X perceived an increase in Tardiness (class), but the behavioral data show that he was not Tardy even once from January through May. Moreover, although he perceived a marked decrease in AWOLs the behavioral data indicate that he went from zero AWOLs in March

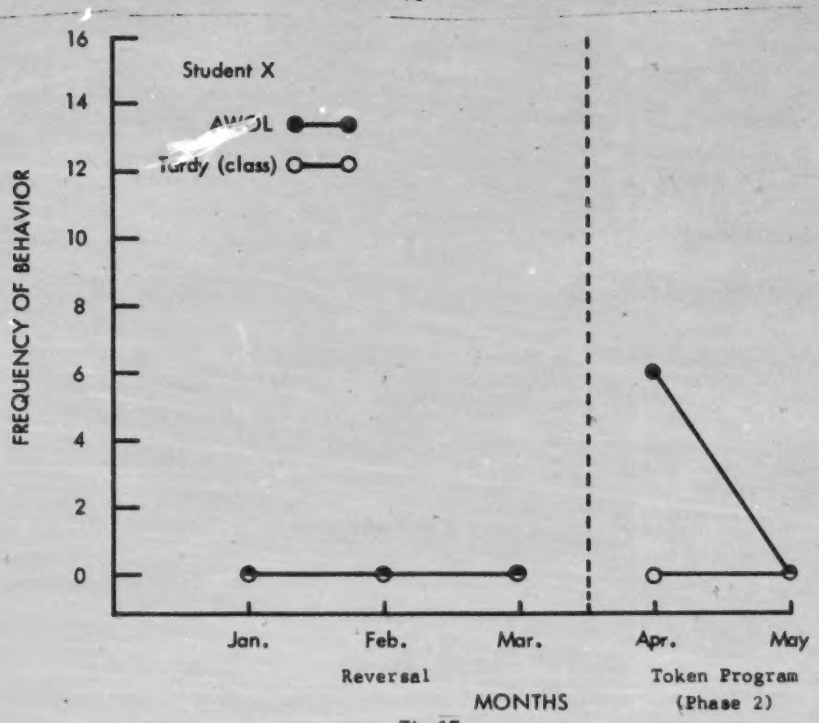

Fig. 15

Frequency of *AWOLs* and *tardiness (class)* for a student perceiving extreme changes in these behaviors following implementation of the token program. Perceived changes: a/Decrease in *AWOLs*, b/Increase in *tardiness (class)*.

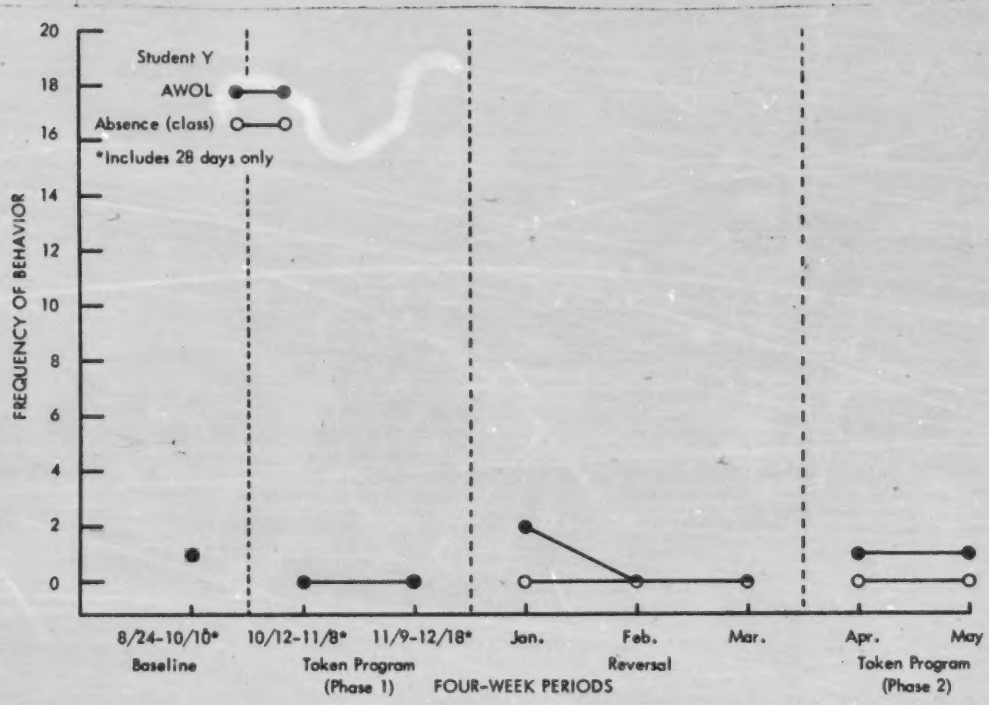

Fig. 16

Frequency of *AWOLs* and *absences (class)* for a student perceiving extreme changes in these behaviors following implementation of the token program. Perceived changes: a/Increase in *AWOLs*; b/Increase in *absences (class)*.

to six in April, the first month of Phase 2 of the program, and then back to zero in May.

Student Y perceived extreme increases in both AWOLs and Absences (class), as illustrated in Figure 16. However, the behavioral records show no absences throughout the first five months of 1971. This student was AWOL once during the Baseline Period, no times during Phase 1, twice during the Reversal Period, and approximately twice during Phase 2. Certainly, these behavioral changes cannot be considered extreme.

As shown in Figure 17, Student Z had no instances of Drinking from late August 1970 through May 1971. However, he "perceived" a marked increase in Drinking during this period. He also perceived a marked increase in Tardy (class), but the behavioral data indicate that he was Tardy eight times from January through May, and only once during Phase 2 of the token program. Thus, for the three students selected, no correspondonce was found between perceptions of behavioral change and the actual changes indicated by the behavioral data.

Additional individual curves were constructed in order to determine whether any behavioral changes were correlated with pretest-posttest change scores on the TSC, ITED, or SRA. In general, no systematic relationships were found between changes in test scores and behavioral changes.

Discussion and conclusions. There are a number of possible reasons for the discrepancies between perceived changes in behavior and the actual changes in the behavioral data: 1) the records may be inaccurate, 2) the students' perceptions may be unrealistic, or 3) replies to the question may not have reflected the students' true perceptions. Unfortunately, there is no way of determining which of these alternatives is correct, on the basis of the available data.

In looking at trends, it was found that there were only a few instances

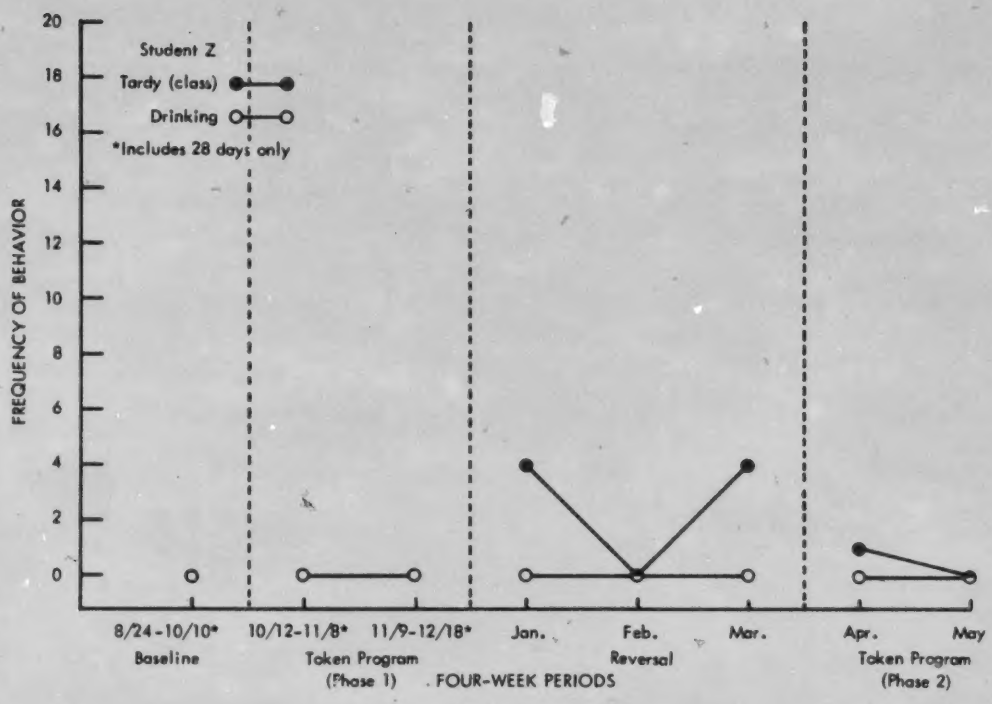

Fig. 17

Frequency of *tardiness (class)* and *drinking* for a student perceiving extreme changes in these behaviors following implementation of the token program. Perceived changes: a/Increase in *tardiness (class)*; b/*increase in drinking.*

where inappropriate behaviors decreased following implementation of the token program in Phase 2 (that is, where the previous trend was either flat or increasing). These were primarily the Absence (class) and Tardy (class) data. The percentage of those Absent (class) for both Kiva males and females increased from February to April, but decreased in May. These changes correspond to the increasing trend in the mean number of absences from February through April, and the marked decrease for Kiva as a whole in May.

With respect to the Tardy (class) data, the percentage of those Tardy increased for Tanoan Hall from January through April, but they showed a decrease from April to May. Frequency data followed a similar pattern. The percentage of students tardy for class increased for Kiva females from January to April, but decreased from April to May for Kiva males. Frequency data for Kiva as a whole also show an increase from January through April and a decrease from April to May.

Thus, on the surface, there appears to be substantial evidence that the MEP did lead to the reduction of Absence (class) and Tardiness (class) both in terms of the percentage of students emitting these behaviors and the frequency of these behaviors for the dormitories as a whole. With a Time-Series Design, however, even when the outcome pattern indicates that the treatment had an effect, there are other plausible hypotheses which can compete as alternative explanations of the shift in the time series. These rival hypotheses arise from the lack of control over extraneous variables which might have influenced behavioral change independent of or possibly in conjunction with the treatment variables.

Probably the most serious weakness of the Time-Series Design is its failure to control for "history," i.e., events other than the treatment variables which may have occurred during the study. For example, suppose

the rate of a certain target behavior after April indicated a change in trend from prior to March. Certain events other than the token program probably occurred during April and May but not during January through March, which may have had some influence on the behavior. A specific instance of this could be the seasonal change from winter to spring which occurred simultaneously with the shift to Phase 2 of the token program. Thus, "spring fever" could account for increases in Drinking and AWOL, precluding improvements which might have resulted had the token program been implemented during a different time of year. Similarly, since examinations are usually administered toward the end of the school year, decreases in Absence (class) from April to May could have resulted from the necessity of having to attend class in order to take final exams.

Furthermore, what appeared to be actual behavioral changes following implementation of the token program, may have resulted instead from a change in the measurement procedure. For example, when the token program was reinstated in late March, the dormitory aides were reminded about keeping accurate records. Thus, the data indicating increases in Drinking may reflect in part more careful observation and recording of this behavior by the dormitory staff.

Another rival hypothesis of behavioral change arises from the fact that the staff knew the predicted outcomes of the study. As Campbell and Stanley (1963) suggest, "If the measurement procedure involves the judgments of human observers who are aware of the experimental plan, pseudo confirmation of the hypothesis can occur as a result of the observer's expectations (p. 41). Or, as in the case of MEP, the hostility of the staff toward the administration and the token program (see Moellenberg) may have produced pseudo "disconfirmation" of the hypotheses.

In conclusion, the reader should remember that the lack of control over extraneous variables in this study contributed to the difficulty of interpreting the results. It is the author's contention that any conclusions regarding causal relationships between the Motivational Environment Program and behavioral changes must be considered tenuous.

Recommendations

The following section contains a series of suggestions which might be of use in subsequent phases of the MEP as well as in similar programs which may be in the initial stages of planning.

1. One of the most important recommendations to be made concerns the selection of an "appropriate" design for the implementation of the MEP, i.e., one which is capable of reflecting causal relationships between the token economy and behavior changes, where such relationships exist. Certainly the reversal design discussed earlier would be appropriate. However, once the program is operating successfully, it may be difficult to justify ethically the temporary termination of the program solely for the purpose of providing objective support for its effectiveness.

A more acceptable procedure would be to employ what is known as the "multiple baseline" technique. With this type of design one would first obtain initial baseline measures on all target behaviors and then begin modifying one (possibly two) behavior(s) while continuing baseline measurement on all the others. Following a clear indication of improvement in the first behavior, one would begin working on a second behavior, continuing with this procedure of successive modification for the remainder of the target behaviors. If changes occur for behaviors which are being treated, while those yet to be treated remain at the baseline level, one would have evidence that the reinforcement variables were indeed influencing the target behaviors.

An important byproduct of this technique is that the particular sequence of target behaviors can vary for different individuals. In other words, one

could work on drinking first for some students, AWOL first for other students, etc. A major advantage of the multiple baseline technique is that the student does not have to try to modify many different behaviors simultaneously. By concentrating on one behavior at a time he is more likely to be successful, and thus obtain the desired reinforcement. Another advantage of this technique is the relative ease with which one can chart the progress of behavioral change. With increased ease of observation and recording, staff members would be more likely to dispense reinforcement consistently and immediately.

Another appropriate design would involve the use of a control group which would receive neither tokens nor back-up reinforcers. Such a group would provide a baseline for assessing the relative effect on performance of the reinforcement variables. Since a number of the students will not participate in the MEP during the next phase, a control group could be established. Of course, it would be best if students could be randomly assigned to the experimental and control groups. If this is not possible, students making up the two groups should be matched on factors such as the baseline rate of the target behaviors. Actually, a combination of the control group and multiple baseline designs would be even more appropriate. However, with a control group design, as with the reversal technique, an ethical problem arises in that the students in the control group would not be given a chance to earn tokens and reinforcers.

2. One of the most deleterious aspects of this study was the system of recording behaviors. Obviously, a study involving as many students and staff as the MEP requires a relatively uncomplicated, standardized record-keeping system to ensure reliable and valid assessment of behavioral change. A recent article by Lehrer, Schiff, and Kris (1970) describes just such a system. These investigators have made use of a credit card in their token

economy, which according to them facilitates the ease of record keeping and concomitantly provides better records for subsequent evaluation.

Under this system the student uses a small plastic credit card similar to a gasoline credit card, on which identification data such as name, I.D. number, etc., can be embossed. Tokens consist of points earned which are stamped on a "points slip." Those who dispense points have code sheets containing numbers which represent specific behaviors. When a specific behavior is reinforced, the identification data is stamped on the points slip along with the number of points earned and the code number. In addition, a space is provided on the points slip for further description of the behavior if it is necessary. The school keeps one copy of these slips, and the student keeps the other. Thus, the student obtains immediate secondary reinforcement in the form of a receipt.

At the end of each week a secretary types out a statement for each student indicating the dates on which he has emitted certain behaviors, the code numbers for those behaviors, the number of points earned, and the student's expenditures for that week. As Lehrer, et al. suggest, the weekly statements permit easy review of a student's progress. Moreover, this technique facilitates statistical analysis of the points accumulated.

3. A third recommendation concerns the different types of behaviors which the MEP was designed to influence. It is likely that behaviors such as drinking and stealing may necessitate a somewhat different approach than behaviors such as tardy for bedcheck and faking illness. For example, the boys in Tanoan Hall received four blue tokens (valued at a total of $2.00) along with verbal praise for not drinking. It is reasonable to assume that for many of these boys the pleasure obtained from getting drunk one night far outweighed the "reward" consisting of $2.00 and a "very good" for staying sober.

The rewards resulting from behaviors such as drinking and stealing are quite complex and may differ extensively for different individuals. Thus, before these types of behaviors are selected for treatment in the next phase of the program, an assessment of the reinforcers maintaining these behaviors for each individual might be helpful. Following this analysis, other procedures for reducing the strength of undesirable behaviors might be decided upon, such as weakening the underlying motivation, withdrawing the reinforcer, etc.

4. The Motivational Environment Program utilized "tokens" consisting of various colored poker chips to bridge the gap between emission of appropriate behavior and delivery of the reinforcers. In various interviews with the staff of Tanoan Hall, the author learned that many of the older boys considered the poker chips to be rather ridiculous. A token does not have to come in the form of a poker chip. For something to constitute a token, it must be: 1) an observable or tangible object, 2) capable of being dispensed immediately following emission of the appropriate behavior, and 3) backed up by effective reinforcers. Thus, poker chips, stars, checkmarks, scores, and a variety of other objects and symbols can be used as tokens. In the future, it may be more beneficial for one kind of token to be used in the dormitory setting and another kind in the classroom. Furthermore, certain kinds of tokens might be better for younger children, such as stars and poker chips, whereas checkmarks might be sufficient for older students.

5. Reinforcers can be determined in a variety of ways. Birnbrauer, Burchard, and Burchard (1970) suggest a number of techniques in addition to token systems which may be helpful in selecting or establishing effective

reinforcers. Three of these suggestions may be of some future value to the personnel at the Albuquerque Indian School. First, and most obvious, is that one could ask the student directly what he likes or wants. However, one problem with this technique is that what a student says he likes may not actually function to reinforce his behavior.

A second method involves the use of the "Premack Principle," which states that if a person is more likely to engage in behavior A than in behavior B when both are permitted, then the opportunity to engage in behavior A will function as a reinforcer for behavior B if A is made contingent upon B. For example, suppose child X is more likely to play baseball than read a book when given a choice, whereas child Y is more likely to read a book than play baseball. In applying the Premack Principle, one would say to child X, "If you read for awhile then you can go outside and play baseball." But to child Y, one would say, "If you play baseball for awhile then you can come inside and read." Of course, this is the type of thing that parents do all the time. Yet, the apparent simplicity of this procedure is deceiving. To be successful in applying the Premack Principle, one must not only be an excellent observer of naturally occurring behaviors, but must also be able to recognize the often subtle cues which would suggest making use of certain high-probability behaviors. Tharp and Wetzel (1969) present an interesting account of the application of the Premack Principle:

> A nurse observed that a slovenly schizophrenic patient, Patricia, invariably engaged all staff members who entered the dayroom in conversation concerning her illegal confinement, together with urging that her lawyer be contacted. The nurse correctly hypothesized this behavior was of sufficiently high probability to be used as a reinforcer; later, opportunity to speak with the staff about her lawyer was placed contingent on hair-combing and face-washing (p. 192).

As Birnbrauer, et al. suggest, however, the Premack technique should not be used exclusively. This is because the opportunity to engage in certain

behaviors may not exist in the natural setting, and thus could not be determined from observation alone.

A third technique involves the use of adult attention, which has been shown to be a powerful reinforcer. Of course, one must be careful to use it in a controlled manner, i.e., making attention contingent upon the emission of desired behaviors and withdrawing attention following the emission of undesirable behaviors.

6. In some situations, reinforcement should be given only when all members of the group emit the appropriate behavior or at least when each member does his part. For example, students should receive reward for cleaning their room only if everyone in the room has successfully completed his particular job. Similarly, it might be better to reward students for being on time to class only when the entire class is on time. At first glance it may appear that this procedure would have an adverse effect on the students who do not successfully do their part, through ridicule from their peers. However, there is evidence that under group contingencies, students seek to help their less successful peers. (Bronfenbrenner, 1962; Wolf, Hanley, King, Lachowicz, & Giles, 1970; Zimmerman, Zimmerman & Russell, 1969). Clearly, with this type of procedure, everyone can benefit only by ensuring that each individual does his part. Thus, to the extent that rewards are dispensed appropriately, i.e., made contingent upon successful performance of the entire group, this procedure should automatically increase helping behaviors from most of the students.

7. Inherent in a good token economy is the selection of appropriate mediators, i.e., those people who are directly involved with the students and dispense the rewards. As noted in Moellenberg's section of this evaluation, a number of staff members were openly hostile to the Motivational

Environment Program and those who initiated it. Obviously, then, one solution is to select those staff members who feel comfortable with the program. However, this procedure would not necessarily ensure successful implementation of the program. Tharp and Wetzel (1969) present the following case which illustrates nicely the importance of the interplay between the selection of reinforcers and mediators:

> A dormitory supervisor, for a boarding school of the Bureau of Indian Affairs, determined that there were two potentially satisfactory reinforcers available in managing the case of a 13-year-old lazy and belligerent boy. The first was probably the most powerful: time with the dormitory-owned electric guitar and amplifier. The mediator for the guitar would be a night-shift dorm aide, Walter. The other potential reinforcer was a town-pass, which would involve a ride in on Saturday morning with the mediator, Billy, the stationwagon driver. The supervisor knew that Walter was a capricious man, whose bad moods sometimes led him to lock up all the entertainment-center equipment: guitars, the stereo, and the pool cues. Billy was a steady young man, interested in "wayward youth," but not at all comfortable with the proposed intervention plan, which seemed to him like bribery. The supervisor believed, nevertheless, that he could influence Billy to dispense rides-to-town on contingency, whether or not there was philosophical agreement. Walter, on the other hand, would require supervision during hours when the interventionist was often out of the building. The mediator of choice was Billy. This eliminated the guitar from consideration, and the plan moved to the second-ranked reinforcer, town-passes (p. 193).

Thus, in this case, the reinforcer to be used was determined by the choice of the mediator. Certainly the intricacies of the reinforcement process require such an analysis before selecting reinforcers and mediators in order to increase the likelihood of a successful, behavior modification program. As Kuypers, Becker, and O'Leary (1968) indicate:

> Great care should be exercised in selecting and training observers, in providing guidelines for the supervisory staff, and in preparing the teacher for what is coming. . . A token system is not a magical procedure to be applied in a mechanical way. It is simply one tool within a larger set of tools available to the teacher concerned with improving the behavior of children. The full set of equipment is needed to do the job right (p. 108).

References

Birnbrauer, J. S., Burchard, J. D., & Burchard, S. N. Wanted: Behavior analysts. In R. H. Bradfield (Ed.), Behavior modification: The human effort. San Rafael, Calif.: Dimensions Publishing Co., 1970.

Bronfenbrenner, U. Soviet methods of character education: Some implications for research. American Psychologist, 1962, 17, 550-564.

Campbell, D. T., & Stanley, J. C. Experimental and quasi-experimental designs for research. Chicago: Rand McNally, 1963.

Kuypers, D. S., Becker, W. C., & O'Leary, K. D. How to make a token system fail. Exceptional Children, 1968, 35, 101-109.

Lehrer, P., Schiff, L., & Kris, A. The use of a credit card in a token economy. Journal of Applied Behavior Analysis, 1970, 3, 289-291.

Tharp, R. G., & Wetzel, R. J. Behavior modification in the natural environment. New York: Academic Press, 1969.

Wolf, M. M., Hanley, E. L., King, L. A., Lachowicz, J., & Giles, D. K. The time-game: A variable interval contingency for the management of out-of-seat behavior. Exceptional Children, 1970, 37, 113-117.

Zimmerman, E. H., Zimmerman, J., & Russell, C. D. Differential effects of token reinforcement on instruction-following behavior in retarded students instructed as a group. Journal of Applied Behavior Analysis, 1969, 2, 101-112.

Section IV

Assessment of Student Attitudes and Perceptions

Marcia Landau, M.S.

Assessment of Student Attitudes and Perceptions

This section will discuss the results of a one to one interview performed at the Albuquerque Indian School. The purpose of the interview was to provide the evaluation team with the subjective reactions of the Indian School student body to the token program. It was felt that nonobjective, open-ended questions would provide a personal depth dimension to the understanding of the effects of the token system on the school. Possibly the students' reports would supplement the "hard data" with subtle areas where the token program was weak or strong. Likewise student recommendations might provide fruitful directions for further token programs in the school.

The interview was performed by 10 Indian college students on a sample of 63 students. Great effort was made to select subjects for the interviews who represented proportionately various subgroups at the Indian School. Age, sex, dormitory, tribal affiliation and academic school attended were factors considered when choosing the student sample. However, a series of extraneous factors, such as student dropouts, student illness, and absence interfered with the acquisition of the ideal sample. There was also a large portion of the school that returned home for summer vacation while the interviews were performed. Students attending the Albuquerque school system had almost 2 weeks more of academic classes than did those at the AIS school. The result is a sample which is somewhat skewed in makeup.

The composition of the sample was:

Sex:

Male :	31
Female:	32

By School:

Attending AIS school:	25	
Albuquerque Public Schools:	38	(These students are referred to as Bordertown students)

By Dormitory:

Dormitories

Wauneka	18
Tanoan	24
Kiva	21

By Tribe:

Navajo	41
Apache	8
Pueblo	11
Utes	3

The interview consisted of approximately 40 questions and a brief multiple choice sheet in which students rated themselves on the problems typically described by the school authorities. The interviews took from 45 minutes to 1½ hours depending on the pace of the individual doing the interviewing and the student involved.

All the interviewers were students at the University of New Mexico. They were all of Indian descent. They were paid on the basis of $2.50 per hour for their services and were trained for approximately 4 hours to administer the interviews. They individually went to the AIS after school hours and located their subjects in the dormitory. They conducted the interviews in semi-private areas and worked at their own pace during the afternoons, evenings and weekends. One hundred subjects were chosen from the rosters of the Indian school and were given code numbers. The identity of each subject was carefully protected by the interviewers and evaluation staff. Confidentiality was focused upon at the initiation of each interview and students were not forced to participate. At the end of a three-week interview period, 63 of

the initial subject list of 100 were interviewed. The interviewing then ceased as the remainder of the students returned to their homes for summer vacation.

The results of the interviews will be divided into five sections for clearer consideration. The first will deal with the students' general attitudes about the Indian school, its personnel and program. The second will deal with the students experiences with the token program. The third will describe the student recommendations for future programs utilizing tokens. The fourth section will report on the students' perceptions of what problems they have most frequently and how the tokens affected the incidence of these problems: Each section will end with a summary and recommendations. The final section will attempt to integrate the results of the interview in relation to the token program as a whole. The data for each section will be listed by question. For each question the number of students responding will be given. The percentages of each response were calculated from the total responding to that question. The number responding varies from 63 to 33 depending on the question. The students' responses which were lacking are due to 1) certain questions which were omitted because the student attended the Albuquerque Public School System which did not use tokens 2) errors on the part of the interviewer 3) certain questions being omitted because the student was too young as judged by the interviewer, and 4) refusals to respond on the part of the student.

General evaluation of school. The section in the questionnaire entitled "General Evaluation" included 10 questions which aimed at eliciting the students' general attitudes toward the school. The questions and response categories and the percentage of students responding in each category appear at the end of this section. The first question asked the students to describe

their attitudes about coming to the school. The answers were distributed such that about 64% of the students replied that they either liked or wanted to come. This is a strong and surprising majority considering the frequent remark by staff personnel that most students are "shipped" there because of behavior problems. Nineteen percent reported some anxiety about the school and 6% frankly stated they did not want to come. Three percent said they were sent by their family and about 5% said they felt nothing. These answers imply that most of the Indian students look back upon their arrival with positive feelings and a minority with a memory of some anxiety. Although we have no comparable non-Indian control group data, this seems to paint a positive picture of the students' anticipation of school.

The next question asked the students to evaluate the school at present. This group of responses again reflects a positive viewpoint with approximately 58% of the respondents reporting unqualified "liking" of the school. Another 26% reported acceptance or positive attitudes with some limitations. This suggests that 84% ranged from neutral to extremely positive reports on the school in general. Conversely there were 8% who reported disliking the school or hating it. Again we see a picture of a student body who in the majority were satisfied with their surroundings.

The next two questions asked more specifically what is "good," or "bad" about the school to provide further understanding of what had motivated the students' previous responses. There was an emphasis on activities and socializing in 65% of the responses to "what is good about the school." The basic task of learning to socialize among peers is appropriate from preadolescence through adolescence and the school seems to be satisfying this function for many of its students. There is a lesser focus (17%) on the positive value of the academic components of the school but this may reflect typical attitudes

of this age group or cultural variations in the importance of school during these formative years. Eight percent remarked upon the new freedoms (getting away from home and liquor privileges) and 5% felt nothing was good. We can cautiously draw some conclusions from these reports. The students are most interested in socializing and learning to meet new peers. Any social functions which the school holds that accomplishes these goals are appreciated. We are also beginning to detect a percentage of students (between 6 and 10%) who are unhappy at the school and resent coming, dislike the school and find little good [illegible] it.

The next question dealt with perceptions of the bad things about being a student at the school. Five general areas of complaint emerged. The largest percentage (35%) complained about the anti-social behavior (fighting, thievery, AWOL and drunkenness). Apparently these behaviors interfered with their routine as well as that of the staff. Thirteen percent complained about the food and 18% complained about the dormitory regulations. Thirty-two percent criticized the school and dormitory personnel. This question is of interest because it reflects the basic problem of this school and any institution dealing with growing and experimenting adolescents. There was reported a need for more freedom and fewer regulations alongside of a report of dissatisfaction and unhappiness with the students who break rules and cause trouble. Perhaps here we see the suggestion of a need for different rules and regulations for those who can use freedom wisely and grow without challenging all norms and separate rules for others who need more structure and assistance until they develop their own controls.

The next question explored the students' general attitudes about their teachers. Here we see a very high percentage (85%) reporting that they liked their teachers, they taught well and that they are "all right." Nine percent remarked that some are nice and again we find 7% either did not like their

teachers or did not care. We see again as in the evaluation of the AIS in general most students like their teachers.

The next two questions tried to again focus on those qualities in the teachers that please the students or make them dislike their teachers. In the area of what they liked about their favorite teachers, it was found that approximately 60% of the students reported that they admired teachers who were nice, friendly, understanding, concerned, kind and those who treated students well. The emphasis in these responses was on the interpersonal quality of the teacher, not knowledge, information, or teaching skill. Twenty-six percent reported that their favorite teachers taught well and a small percentage gave scattered reports about teachers being in control, looking well, not giving homework or looking well groomed. We may conclude that the majority of these students are primarily social beings who respond to adults on the basis of how they are treated. Warmth and friendliness are strong assets. We also see emerging a repetitious percentage of 10 to 15% of the students who are academically oriented and focus on the learning aspects of the AIS and academic work. Approximately 8% of the answers seem to reflect neither social nor academic awareness (looks well groomed and not much homework). These would be hard students to keep happy under any circumstances.

When asked to report on what they liked least about their worst teacher a wide variety of responses was obtained. Thirty-three percent reported they disliked teachers who were irritable, mean, nasty or rejecting. Fifteen percent resented poor teaching techniques. Eighteen percent felt the demands made were unreasonable. Three percent complained that teachers were nosy or they did not like their room. The picture here is similar to that described before. Most students were upset by adults who treat them rudely. A small percentage just complained about minutia. There was also a noticeable percentage

(21%) who reported all teachers are good. This is the first instance where we see a hesitancy of some students to constructively evaluate, criticize or complain about their environment. This may be due to concern about confidentiality (although this was discussed before each interview) or a sense of overwhelming humility in the face of academic personnel who are frequently Anglo and successful participants of the higher institutions. The answers suggest that many students need teachers with much personal patience and sensitivity. Others need teachers who teach at a cautious pace, aware of students who are struggling and some need teachers to encourage much development of verbal judgment and opinion formation on the part of the student.

The next questions focused on general and specific views of the dormitory personnel and the students responded with less enthusiasm than when they described their teachers. Around 30% were described as nice or generally nice. Forty-seven percent received descriptions which were qualified. These ranged from mostly nice to 50-50 to mostly bad. Fifteen percent had strong criticism including nosy, irritable, bossy, untrusting, too punitive and angry. The students' ability to criticize the dormitory people more freely and frequently may be due to significant personality differences between dormitory and academic personnel, or the greater degree of closeness between students and dorm personnel due to the time they spend together or the social closeness between the students and dorm personnel who are frequently of Indian heritage. The complaints, although more vociferous, are still reflecting the students' feelings that they found adults treated them in an authoritative, hostile fashion and they resented this when it happened.

When the students were asked to describe the qualities of their favorite dormitory person, the "humanitarian" theme dominated descriptions. Almost 90% of the descriptions included behaviors on the part of the dormitory person in which the student was treated like a friendly peer rather than a younger

and therefore incompetent person. They liked personal contacts, help with personal problems, patience, someone who joked with them, kindness and someone who stretched regulations under special circumstances. Thirteen percent evaded the answer, as they did when asked to describe their teachers' faults. They reported all dorm people were the same. Only 2% were unable to come up with any kind words for any dorm person and felt they were all mean or awful.

When asked about the qualities of the dorm people they disliked, 19% evaded the questions saying all were the same or "I don't know." Another 14% said nobody had bad qualities. Here we see again another reflection of the tendency of some students to avoid open evaluation of their superior even in a situation when they are assured of confidentiality. Sixty-eight percent complained about dorm personnel who were short-tempered, pushy, who punished severely and frequently, who used physical punishment, who were untrusting and at times ignored important student requests. This may be another reflection of the Indian students' resentment of what they interpret as supercilious and overly authoritarian treatment.

Summary:

These general questions provide the foundation for understanding the mood of the Indian students and the AIS atmosphere. In general, this seems to be a heterogenous group of students. Most enjoyed the school and dorm and were especially involved in learning social mastery with their peers. They liked their teachers and enjoyed close student-adult relationships and mature treatment. They resented being ordered about or "talked down to" and usually evaluated their teachers on the basis of their social skills rather than their knowledge or teaching ability.

They were less pleased with the dormitory and dorm personnel but still seemed to accept the school life because of the social activities. Some were

afraid to speak out and others complained no matter where they were. At this point, it appears that individualized programs of reinforcement are appropriate as the students differ greatly in where they lie in the realm of maturity.

Student Questionnaire

I. General Evaluation of the Albuquerque Indian School.

1. Can you remember and describe how you felt about coming to this school? (N = 63)

Percent Response	Responses
34.9	like to come to school
28.6	want to come to school
14.3	scared or confused
4.8	lonely or afraid
4.8	don't want to come to school
3.1	don't like to come to school
3.1	sent by family
1.6	no comment
4.8	felt nothing

2. What do you think of the school in general? (N = 62)

Percent Response	Responses
58.0	good or likes
14.6	O.K. or alright
11.3	likes but qualifies with dorm
6.4	likes a lot--great
6.4	don't like
1.6	hates
1.6	no comment or don't know

3. What are the good things about being a student here? (N = 60) (40 students, with 20 giving 2 responses)

Percent Response	Responses
31.4	activities in general
22.5	friends relatives--socializing
17.6	school
11.3	dorm activities
8.8	getting away from home--freedom
5.0	nothing
1.2	drinking privileges
1.2	everything

4. What are the bad things about being a student here? (N = 76) (52 students with 24 giving 2 responses)

Percent Response	Responses
21.0	fights and thefts, AWOL
18.4	dormitory life (regulation)
14.4	drunkenness
13.1	food
11.8	nothing

11.8	school
10.5	dorm personnel
3.9	leave regulations

5. What do you think of your teachers? (N = 46)

Percent Response	Responses
47.8	like them, nice
21.8	O.K., alright
15.3	teach good
8.6	some nice
4.4	don't like
2.1	nothing, don't care

6. What do you like most about your favorite teacher? (N = 38)

Percent Response	Responses
44.7	nice, no explanation
26.3	teaches well, explains well
7.8	friendly
5.2	understanding, concerned
5.2	not much homework, not strict
2.7	treats students well-kind-gives extra help
2.7	looks well-groomed
2.7	have control of class
2.7	interesting

7. What don't you like about your least favorite teacher? (N = 33)

Percent Response	Responses
33.3	irritable, strict, mean-sh[illegible] tempered, nasty, yells
21.2	like all teachers or none bad
18.2	unrealistic demands
15.2	teaches poorly, expresses self poorly
6.0	boring or repetitious
3.0	rejects students
3.0	nosy, don't like room

8. What are the people who work in the dorm like? (N = 61)

Percent Response	Responses
29.5	all nice, generally nice
22.9	50% nice, 50% bad--or O.K.
13.2	most bad, some nice
11.4	most nice, some bad

8.2	nosy, irritable, hostile, bossy, gossip
8.2	O.K., unless you disobey
3.2	not close enough-warm enough, untrusting
1.7	all bad
1.7	punishment too frequent, too strict

9. What do you like about the dorm person you like best? (N = 62)

Percent Response	Responses
22.6	help w/personal problems or school work, or job placement (understanding)
14.6	gives special privileges (stretches regulations)
14.5	nice
14.5	kind, not bossy, or short-tempered
12.9	does favors-lend money or give ride in car--gives tokens
12.9	all the same (both directions)
3.2	jokes or fools around
3.2	talks to students a lot
1.6	all mean or awful

10. What don't you like about the dorm person you like least? (N = 59)

Percent Response	Responses
44.0	bad mood, easily gets mad, mean, pushes students, cusses a lot
18.7	all the same or I don't know
13.6	nobody
8.5	too fussy about chores
5.0	punishes too much, too often too severely
5.0	uses physical punishment
3.4	doesn't listen, not trusting, jumps to conclusions, gossips
1.7	at times ignores request (mail-washing machines)

Evaluation of token system. The second section deals with the students' reactions to the token program. The questions and response categories and the percentage of students responding in each category appear at the end of this section. Students were first asked how they found out about the program. Fifty-eight percent reported learning about it at a general meeting and 5% reported a meeting with Mr. Blanchard. The rest of the answers varied including learning about the program "at class," "talking to a friend," "I read it on the bulletin board," "I learned about it at a workshop " and through a meeting with the principal. There is therefore a group of 32% who apparently learned of the system in very varied fashions. One would hope that a more controlled standardized indoctrination would be preferred. Perhaps a series of meetings (not one) in the dorm coinciding with coordinated classroom meetings would be the ideal. In this way, no student would be left to learn randomly about the program from friends or from printed material alone. It also might be beneficial to have regular talk sessions each month in which the students provided feedback on the effective and ineffective aspects of the token system.

The second question dealt with whether the student felt anyone had asked him for his opinion about the program. Fifty-two percent responded that no one asked for his opinion. The rest of the students varied in responses. Some reported that dormitory personnel, teacher or friends asked their reactions. Sixteen percent said they filled in a formal questionnaire. Here again we see that from the students' point of view their opinion was not sought in any regular or organized fashion. This may be important as the results obtained from the first section of the questionnaire implies that the students are most cooperative when their opinions are treated with respect and their feelings are perceived by adults. Feedback from students should be sought and made public

in future programs so that there will be a sense of participation and investment by each student in the program.

The next two questions sought out value judgments on the part of the students as to which goal of the program they felt was the most important and which they felt was the least important. Thirty-three percent of the students responding felt their dormitory work detail was the most important goal to be reinforced. The rest of the students were divided into smaller groups who felt that bedcheck, class attendance, not fighting nor getting drunk, and working at jobs were important goals. A small percentage felt homework and punctuality were important. These answers give further clues into the values of the Indian students. Apparently they are more interested in doing their own job or task rather than following behavioral regulations. This reflects an attitude which values a job accomplished rather than behavior for its own sake.

When we examine the reports of the least important tasks, there seems to be a wide distribution of responses. However, a blending of responses suggests that punctuality is a goal which the students do not value. About 58% criticized the need for roll calls, being on time in general, and bedcheck as unimportant goals. This may reflect a clash of the values between the students and the administration in their different focus on the importance of time and punctuality.

The next two questions asked how the students got the majority of their tokens and what caused them to get the smallest number of their tokens. Forty-eight percent of the students reported getting most of their tokens doing their job detail and school work. Fourteen percent evaded the question saying that they got about the same from everyone. Four percent reported getting tokens for extra details. The question of why they got few tokens from some personnel evoked vague and inexplicit answers. The students seemed to know that

at certain times or when around certain persons they got fewer tokens but they were not able to determine what caused this.

When asked if the students felt they always got tokens when they deserved them 51% said they did. Twenty-four percent said they did not but were unable to explain why. There were a few specific complaints about instances of unfair judgment, absent-minded personnel and the supply of tokens running out.

Thus it can be seen that most students perceived the tokens as distributed fairly. Most got them for work at school or in the dorm. Most disliked the focus on punctuality in the program. The students did not feel, in general, that they helped design the program and most found out about the program in a group meeting.

The next set of questions asked the student to judge the effectiveness of the token system at the AIS. When asked whether the program was good or not good for the school, 85% of those responding reported it was a good program. Six percent felt it was both good and bad and 8% felt it was bad. The responses to this question certainly present a positive student endorsement of the token system. However if we examine the "why" behind the students' approval we see that only 13% felt it improved behavior or learning. The rest of the students supported it because it gave them more money and privileges. This lack of focus on the students' part on behavioral consequences may have reflected the results of no feedback to the students of the goals of the program and its efficiency. Perhaps if adults and students planned together on what behaviors needed change and how to utilize the tokens and then followed the behavioral results of the program each month, their view would be less opportunistic and more behavior oriented.

When asked if they purchased anything they really wanted, 53% replied

yes. Twenty-nine percent said no but offered no explanation and 18% explained that lack of the right merchandise, a shortage of items in the store and the utilization of the tokens too quickly prevented them from getting what they wanted.

The ninth question investigated the saving and spending habits of the students. The responses suggested highly varied saving patterns among the students. A few students saved for 2 or 3 months. Most saved for a week to a month. Fourteen percent spent their money right away and a small number complained about having tokens stolen. It appears that realistic planning for the students' spending habits should be based on items which can be purchased within a month's time. Items or privileges which take longer will only be available to a small percentage (7%) of students.

The students were asked if there were things they wanted but could not purchase with the tokens. Thirty-six percent said no and the remainder were divided among unsatiated desires for food, T.V.'s, more clothing, field trips, stamps and vacation trips. These all seem to be feasible possibilities or additions in a future token program. It is interesting to note the importance of leisure and leisure activities in this list. Apparently the students liked to plan their leisure and are willing to work for new leisure possibilities.

The next question asked whether the students felt dormitory or teaching personnel tried to control their use of the tokens. Only 17% said yes without giving any explanation. Thirty-eight percent said no. The rest of the respondents said yes, they were controlled because the attendants got them to do things (work, not drink, get to class, etc.). This question was aimed at discovering whether the students felt manipulated by the token system. Apparently they did not feel personally manipulated but they did realize they were conforming in some areas because of the token reward.

The next five questions were designed to determine how much the students knew about the administration and planning of the token program. When asked whose idea the program was, 78% students replied they did not know. Twenty-one percent thought Mr. Blanchard designed it and 2% thought it was the attendants idea. These responses suggest a surprising ignorance on the origins of this program which probably leaves room for more fancy than fact. It is recommended that the source of future programs be stated clearly to the students so they do not feel that vague administrative author'ties are indirectly controlling their education.

The confusion was further evidenced when the students were asked why the program was started. Thirty-eight percent of those responding did not know. Eighteen percent thought it was designed to increase students' possessions and finances. Forty-four percent related it to student behavior but most of these saw the goals as one specific behavior instead of a total approach.

The next two questions asked how the students thought the behaviors to be changed were chosen and how the number of tokens for each behavior was chosen. Over 50% did not know how these decisions were made. Again we see that the students were not clear on how decisions were made and this possibly interfered with their ability to respect the goals and methods of the program.

When asked who decided what they were able to purchase with tokens, the responses were mixed ranging from don't know, to dorm personnel, to Mr. Blanchard, to token store personnel, the teachers, and the administrators. Some clear messages are needed here.

Summarizing these responses we see that in general the students were not clear as to why the program was begun, who designed it and how the goals and rewards were established. Clearer feedback from the authorities is recommended.

The next two questions asked the student to hypothesize what would occur

if they were disobedient in the dorm or classroom. The purpose of the question was to see if the students reported standardized punishment or handling of these problems. In the dormitory the students were asked to imagine what would happen if they came in drunk. The replies included a large percentage, 56%, who said that the dorm personnel would vary in their response depending on the disruptiveness of the student. Punishment could range from being sent to their room to token penalties or to a more severe penalty of being sent to the D-Home. The message seemed to be that drunkeness was a mild offense compared to surliness and hostility on the part of the student and the latter was more severely punished.

When asked what would happen if they disobediently spoke in class the responses were greatly mixed. They included being sent to office, having a grade lowered, a token penalty, getting yelled at, getting slapped or being reported to dormitory. The comparative responses on these two questions suggest that the dormitory personnel were more consistent in their response to the suggested crises, whereas teachers varied greatly from class to class.

Summary:

This section reviewed the students' reactions, evaluations and impressions of the token program. In general we can see that students were not clear about the origins or purpose of the program. They learned about it from different sources and felt their opinion was not important in its planning. They liked the program, especially receiving tokens for doing work in the dorm or at school. They resented the administration's focus on punctuality and were less appreciative of tokens received for this kind of conforming behavior. They felt they got tokens when they deserved them and liked the program for the school. They bought items they wanted and saved up to a month at times for a special purchase. They would have liked to

have had a broader range of activities, privileges and items available for tokens. They found dorm personnel more predictable in their handling of misbehavior than teachers and placed little focus on the use of tokens to handle serious misbehavior.

Recommendations include providing the student with much more information on the origin and planning of government sponsored programs. Student reactions and feedback should be sought and incorporated into the system and organized channels should be provided for griping or recommendations. A much more extensive training program is probably required to help the student to understand the connection between the tokens and their behavior.

The tokens might also be used to purchase trips, privileges and activities according to the students' wishes. Also students might determine the financial value to be placed on each behavior. Overall, student participation, feedback, and training, are recommended.

Student Questionnaire

II. Reactions to Token Program

1. How did you first find out about the token program? (N = 62)

Percent Response	Responses
58.0	in dorm meeting (at assembly-matrons)
12.9	in class at school
9.6	talking to friends
4.8	don't know
4.8	meeting--Mr. Blanchard
3.3	meeting--w/principal
3.3	work shop, home ec
1.6	bulletin board or memorandum
1.6	at store

2. Did anyone ask for your opinion on the program? (N = 61)

Percent Response	Responses
51.6	no one
16.1	yes, filled out questionnaire
13.2	yes, dorm people
4.9	yes, in discussion
3.3	yes, teachers
3.3	only friends
3.3	Mr. Blanchard
3.3	I don't know or can't remember

3a. Which are the most important tasks for which you receive tokens? (N = 48)

Percent Response	Responses
33.3	doing dorm detail
11.9	bedcheck
11.9	all
9.6	going to class
9.6	not fighting (behavior)
9.6	not getting drunk
7.1	working at job
4.7	punctuality
2.3	homework, classwork

3b. Which are the least important tasks for which you receive tokens? (N = 36)

Percent Response	Responses
22.2	roll calls
19.4	being on time (bus)
16.6	bedcheck
13.8	room clean or details

8.4	getting drunk
5.6	extra details
5.6	AWOL
5.6	all unimportant
2.8	school attendance

4. Who gives the most tokens? Why? (N = 49)

Percent Response	Responses
24.4	doing job details
23.9	teachers & good school work (just give more)
14.2	all same
14.2	no explanation given
10.2	dorms give more
8.1	being good
4.0	gives extra details
4.0	don't know

5. Who gives you the least tokens? Why? (N = 47)

Percent Response	Responses
23.5	don't know, not clear
19.2	dorm people or person
17.0	I don't do required task
14.8	school people or person
10.6	all give the same or none
8.5	mean, doesn't get along
6.3	person too busy or forgets often or new personnel

6. Do you always get tokens when you think you should get them? If not, how come? (N = 49)

Percent Response	Responses
51.0	yes
24.4	no, don't know why
8.1	sometimes
4.1	don't know
4.1	unfair judgment, I disagree
4.1	person forgets
2.1	not clear answer
2.1	person runs out of tokens

7. Do you think the token program is good for this school? Why/Why not? (N = 48)

Percent Response	Responses
58.3	good--get things we need
12.5	behavior is better, learning is better
10.4	helps poorer students

6.2	good and bad
4.2	wrong for school students, inappropriate
2.1	good for superficial
2.1	makes student feel like guinea pig
2.1	wrong to pay for good behavior
2.1	helps students

8. Have you gotten anything you really wanted? (N = 45)

Percent Response	Responses
53.3	yes
28.9	no, no explanation
8.9	no, use tokens for other items
6.6	no, they don't have what I want
2.3	no, what I want is gone--sold

9. If so, did you save for it? How long? (N = 42)

Percent Response	Responses
30.9	up to a month
16.8	didn't want anything
14.3	up to a week
14.3	spent right away, took too long
11.9	don't know or can't remember
4.9	up to two months
2.3	up to three months
2.3	money stolen
2.3	not too ambitious, don't get enough

10. Is there any privilege or thing you want a lot but you can't get with tokens? (N = 42)

Percent Response	Responses
35.8	no
14.3	food
11.9	T.V. or radio
11.9	clothes
11.9	field trip or movies
9.6	home trip or other vacation trip
2.3	stamps
2.3	money

11. Do you feel the teachers or dorm people try to control you using the token economy? (N = 47)

Percent Response	Responses
38.2	no
23.4	yes, to get us to work
17.0	yes

6.4	yes, to make us behave
4.2	yes, to get us not to drink or AWOL
2.2	sometimes
2.2	yes, to get us to roll call
2.2	yes, to get us to class

12. Whose idea was it to start this program? (N = 49)

Percent Response	Responses
77.5	don't know
20.5	Blanchard
2.0	aides

13. Why did they start it? (N = 45)

Percent Response	Responses
37.8	don't know, none
15.6	good to increase student (activities, possessions)
13.4	good for behavior, help student (vague)
8.8	to make us behave
6.7	to control behavior
6.7	to get us to do work or detail
4.4	to get us to go to class
4.4	good to keep students out of trouble
2.2	to help poor kids

14. How did they pick the behavior to earn tokens? (N = 44)

Percent Response	Responses
50.0	don't know
20.4	chose negative behaviors
9.1	chose important behaviors
6.8	chose job rewarding behaviors
4.5	asked the dorm personnel for recommendations
2.3	chose behaviors stressing individual accomplishments
2.3	chose negative behaviors for bad kids only
2.3	asked the students to decide
2.3	misunderstood question

15. How do you think they picked how many tokens for each behavior? (N = 48)

Percent Response	Responses
52.1	don't know

20.8	misunderstood question
12.5	gave more tokens for bigger problems or harder tasks
6.2	student council or body
2.1	decided administration and then listened
2.1	staff
2.1	had limited supply and ran out
2.1	no comment

16. Who decided what you were able to get with the tokens? (N = 43)

Percent Response	Responses
32.6	don't know
16.3	dorm personnel
13.9	individual being rewarded
13.9	Mr. Blanchard
9.4	students
6.9	token store personnel
4.6	superintendent--administrators
2.4	school personnel (teachers, principals)

17. What would happen if you came in drunk to the dormitory? (N = 48)

Percent Response	Responses
56.3	depends, can range from going to bed to D-Home and expulsion
10.4	D-Home and not token
10.4	no tokens for a week or shorter
8.4	restriction (with or without tokens)
6.3	don't know
4.2	sent to room or office
2.0	nothing
2.0	write parents

18. What would happen if you disobeyed a teachers request to be quiet? (N = 33)

Percent Response	Responses
33.3	sent to office
15.1	give lower grade
1.1	no tokens
12.1	yell at you, take away tokens
6.1	depends on the sex of the teacher (talk or slap)
6.1	depends on the teacher
6.1	don't know or misunderstand
3.0	report to matrons

Students' estimates of change. In this section the results of a rating scale on the behavioral goals of the study is reported. Students were asked to rate themselves on the behaviors before and after the token system. There were 5 choices for each item ranging from never (1 pt.) to all the time (5 pts.). The scores before and after on each item were compared and a difference score on each item for each subject was calculated. The results are listed in Table I. The most striking result in the data is the large percentage of subjects who perceived no change taking place on the behavioral items. at its highest, 98% of subjects perceived no change on the item #9, "getting in trouble because of sex", and at its lowest 62% of subjects perceived no change on item #2, "lateness to class". Where change does take place, it seems to go in both directions suggesting that the students have not found the program to be a large asset in improving their behavior. The items with the most change were:

1. Lateness to class
2. missing class
3. missing bedcheck
4. missing dorm meetings
5. misbehaving in class

However the mixed directions of the scores suggest that only "missing dorm meetings" had a strongly higher percentage of subjects decreasing their poor attendance than increasing poor attendance.

Least affected were:

1. getting in trouble because of sex
2. destroying property
3. taking others' property
4. making believe you're sick
5. picked up by police.

There is clearly more fluctuation on the less serious problems than on the more serious problems from the students' perceptions. It is possible that the high variability on the less important items may relate to the students lack of interest and concern with behaviors such as punctuality and roll

call. Here perhaps is the most marked clash between the values of students and administration.

Also calculated were the mean and standard deviations of the students' estimates regarding the current frequencies of each target behavior (Table II). Their perceptions indicate they had the most frequent incidences of the following problems:

1. lateness to class
2. missing class
3. missing dorm meeting
4. AWOL
5. missing bedcheck

It is obvious that 4 of the 5 items on the most frequent list are on the most variable list (most change). These problems are typical problems of adolescents in a dormitory setting and may reflect the passive aggressive and careless way students respond to general institutional rules. Perhaps these problems more than others are in the realm of student awareness and control. They may change more dramatically and appropriately if students choose them as behavioral change goals and plan their own program.

Behaviors with least frequent incidence as perceived by the students were:

1. taking others properties
2. sex trouble
3. picked up by police
4. destroying property
5. making believe you're sick

These are serious anti-social problems but the majority of the students did not seem to feel they applied to them. The individualized responses on this Table (II) give further support for more specialized token programs with students choosing their own goal or goals and checking their own behavior each month. If the students' perceptions are correct, it is a waste of tokens and effort to concentrate on these more serious problems with the majority of the students. Also recommended is double feedback with students assessing

change in their own behavior each month and then comparing their estimates of their change in behavior to that of dorm or teacher reports. This would help students to learn to be more realistic about evaluating themselves.

Student recommendations for future programs. This section focused on the specific recommendations which the students selected for future programs. The first question asked the students who they would have help them plan a hypothesized new token program. The students were divided in their approach but 36% chose a blend between themselves and a more experienced adult. Twenty-one percent said they would use all friends and students. Eleven percent suggested the use of dormitory aides and 3% suggested their teachers. It seems that the students have most confidence in their own skill combined with an adult and the preferred adult seems to be a dormitory person. The student council was also mentioned and may have some beneficial use in future planning. It seems there was little emphasis on the concept of democratic planning (rather than bureaucratic planning) in the students' thinking. The student council may provide future programs with a vehicle for getting at student ideas and teaching the ground rules of democratic prodecure at the same time.

The next question asked what behaviors the students would pick to change in their program. Here we see a large percentage (73%) chose the same list as is now in use or items from the list. The additions are "no backtalk" and more focus on good behavior. These responses reflect the fact that students see the basic adjustment problems in ways which are similar to the administration. Allowing them to come up with some list may have much more meaning in establishing their motivation.

The third question asked if they would use tokens or something other than tokens. Again we see a basic support of the token plan as 68% chose tokens. The other suggestions of awards, tickets, activities, and money may

be employed as things to be purchased with their tokens. Perhaps students might be allowed on an experimental basis to turn in hypothetically 10% of their tokens for money at a low rate of exchange. It would be interesting to see how they would handle such an option.

The fourth question directly focused on what is missing in the reward aspect of the current program. Here we see a reiteration of question 10 of the Token Evaluation section. Students wanted freedom and social activity in exchange for their tokens. This included trips, privileges, food (snacks), T.V. and some suggested the availability of more expensive items like cars or musical instruments. We still see 48% endorsed the current plan and requested no change.

Finally students were asked to describe how they would like to see such a program explained to the students. Sixty-five percent said that they would use the same method. Additions included more personal contact, a handbook, a seeking of student feedback and student efforts in planning and 2% suggested showing the students the store as a motivator. All these suggestions have merit and it seems that planning as careful as that which went into deciding token values is needed to plan how the students enter and particiapte in this program without instilling resentment and evasion.

Summary:

The student recommendation of this section suggests that they are basically satisfied with the program and have a few ideas to add for improvement. They asked to be able to participate in the planning and administration. They wanted to work with the dorm personnel in doing so but placed a great deal of value on their peers' reactions. They accepted the goals of the program but there was a degree of disinterest evident in their responses. Perhaps they have never considered a question of educational goals and

agreeing with the establishment is the quickest way out. They approved of tokens as rewards but wanted to be able to buy more with them. They were especially enthused about being able to go to more activities and have more freedom of movement. They again basically approved of the introduction of the program to the students, but suggested a program supplemented by booklets and feedback.

These are very important questions and one wonders how the answers would compare if students were encouraged to come up with consensual answers after a one hour meeting and discussion. In many cases, there seems to be a lack of interest and motivation on the part of the students to conceptualize and plan for themselves.

Student Questionnaire

IV. Student Recommendations for Future Programs.

1. Who would you have help you plan your program (pretend you are in charge of a new token program)? (N = 61)

Percent Response	Responses
36.0	combination or someone experienced
21.4	friends and students
11.4	dorm aides
9.8	don't know, misunderstand, or no one
6.6	administrators or school board
6.6	Mr. Blanchard or psychologist
3.3	student council
3.3	teachers
1.6	everyone

2. What behaviors do you think you'd pick? Why? (N = 56)

Percent Response	Responses
51.8	selected beahviors from present list
21.5	same as now
10.8	don't know
7.1	no backtalk or meanness
5.3	focus on good beahvior
3.5	not clear

3. Would you use tokens or something else? (N = 59)

Percent Response	Responses
67.8	tokens
18.6	money
3.4	don't know
3.4	tokens and money
3.4	awards
1.7	tickets
1.7	activities

4. If you were using tokens, what kinds of things or privileges do you think the tokens could purchase ? (N = 59)

Percent Response	Responses
47.5	some
16.9	add trips off campus (extra curricular)

6.8	add privileges
6.8	add food
6.8	exchange for money
5.1	some but add more
5.1	I don't know
3.3	cars or T.V.
1.7	add musical instrument

5. How would you get the students to understand the program before it began?

Percent Response	Responses
65.4	some
10.3	more individual effort
6.8	handbook and meeting
5.2	I don't know
5.2	get feedback
3.5	let students help set up
1.8	more explanation
1.8	show students the store

Discussion and Conclusions

A long and extensive questionnaire battery such as the one used to interview students supplies the reader with a great deal of diffuse information. The goal of this section is to seek out patterns and clusters of information which will give us more insight into students' thinking and feelings. Then if we succeed we can see how well the goals of a token program fit into the psychological organization of the students at the school.

There is repeated evidence in the responses of the students that they are not at the Indian School to complete a step or milestone on the way to a "career" in the white middle class sense of the word. Instead, they are at the school to experience life away from home, to learn to make new friends, to learn their role as a youthful member of their respective tribes in relation to other Indians and Anglos, and to experience the social and entertainment benefits of living in a large city in contrast with a small rural environment. There is little future orientation in their thinking, and they instead mark academic time at the school, while their social skills bloom and develop.

As a part of this attitude, which is analagous to a white child's perception of a summer-camp environment, authorities are seen as people who put on the controls and are usually distant in terms of personal identification. Personnel who take time to become involved or intimate with students earn a special place in their lives. Irritable, punitive authorities are stereotyped into the institutional roles as angry parents. When asked about their pleasures and use of the tokens, we observed mainly social and oral concerns. Hobbies and skills were not mentioned nor was there a seeking a specialized knowledge for the future.

Overall this paints a rather bleak picture of the student who is more concerned about his peers than his schooling; more interested in playing than

learning; more involved in rebelling against adult models than in imitating them, and more desirous of outer controls of inappropriate behavior than in developing inner controls. The bleakness arises because the school and its traditions go in direct opposition to these qualities and expect the students to be conscientious, future oriented, and well-controlled.

If we accept the goal of mature responsible behavior and positive self-concepts as a goal for the Indian student, we must ask how does the token program fit into the attainment of these goals. In general tokens reinforce specific behavior, not attitudes. They were applied across the board to all students regardless of individual differences. They had no built-in method of focusing on special strengths or weaknesses of each individual student. They supplied the students with spending money which did allow some degree of choice, be it for specific purchases in terms of whether to save or spend. They were handed out by the authorities, not by the students, and the amounts and focal behaviors were established, according to the student's perceptions, completely by the authorities. In other words, the token system is almost a perfect reflection of the message that tells the student:

1. You cannot control yourself.
2. We will list and remind you as a group of your control weaknesses.
3. We will establish rewards for all who do not do things which reflect lack of control.
4. We will determine how much and how often you will receive these rewards.
5. We will also tell you where to spend the rewards, allowing some leeway and choice.
6. We will be able to measure the amount of control you have acquired by the number of tokens you have or the merchandise you have purchased.

This series of messages can undoubtedly instill resentment of authority and an attitude of opportunism in the students in relation to tokens. The tokens become ways of getting goods and the rejection of tokens adds a new way of rebelling against institutional goals.

The picture is not as negative as it appears. The students ask in their responses for more involvement in planning the program. They ask for more decision-making on the use of the rewards. They ask for more equalitarian treatment by dormitory personnel and teachers. They ask for more focus on work and accomplishment rather than on obedience for its own sake. They request more feedback on how they are doing and information on where decisions concerning their lives are made.

In essence it appears that the token program might have real potential in helping these students to become more involved in their own education and in controlling their behavioral environment. Concepts like cooperation, group decision-making, voting and electing of representatives, are all ideas which encourage people to share views, modify their thinking, listen to others, and learn to voluntarily compromise when it is necessary. Students might be encouraged to utilize such concepts to fashion future programs. Students might be quite resistant to taking on all this planning at first and the enormous responsibility of policing themselves and giving feedback in a diplomatic way to difficult students. However, it is recommended that this method be tried even on a small level with a limited group of students to see if it has encouraging implications for future work. A small percentage of a dorm or a mixed group might be selected to work out with professional guidance their own token system with decisions on behavior, goals, rewards, and the use of rewards to be left to them. Such a trial program might encourage strong feelings of personal worth, competitiveness with other students on "good

behavior" and might give these students their first positive experience in shaping their own destiny. With an awareness that this is highly a controversial and unusual use of the token system as it now stands it is still offered as a tentative and possibly benefical use of a system that seems to have little deep meaning to the students at present.

Finally it is recommended that training become a much more complete intensive, and long term project. The idea of giving a token for a behavior seems simple at first but only a great deal of experience can place the appropriate emphasis on consistency and praise as necessary ingredients for real change. Students (and perhaps staff) might do well to focus on only one behavior at first and experience the complexity of consistent reinforcement so that they can fully understand how a token system operates.

The token system has great potential for helping these students mature but a more student oriented planning, training and evaluation may be necessary to fully realize the benefits of the system.

Appendix B

Table 1

Student Estimate of Change During the Program

Behaviors	Students Reporting Change Number	Percent	Decrease Number	Percent	Increase Number	Percent
1. missing class	23	36.5	6	9.5	17	26.9
2. lateness to class	24	38.1	10	15.8	14	22.2
3. misbehaving in class	17	25.4	13	20.6	5	7.9
4. drinking alcohol	9	14.3	3	4.8	6	9.5
5. AWOL	15	23.8	9	14.3	6	9.5
6. missing bedcheck	20	31.7	11	17.5	9	14.3
7. police pickups	7	11.1	5	7.9	2	3.2
8. destroying property	3	4.3	1	1.6	2	3.2
9. sex trouble	1	1.6	0	0.0	1	1.6
10. missing dorm meeting	20	31.7	16	25.4	4	6.4
11. misbehaving in dorms	9	14.3	5	7.9	4	6.4
12 making believe you're sick	7	11.1	6	9.5	1	1.6
13. taking other people's property	4	6.3	4	6.4	0	0.0

Note.--N = 63.

Table 2

Student Report of Severity of Problem

Behavior	Mean	S.D.
1. missing class	2.3	.95
2. lateness to class	2.4	1.2
3. misbehaving in class	1.5	1.2
4. drinking alcohol	1.8	.9
5. AWOL	2.1	1.2
6. missing bedcheck	2.1	1.5
7. police pickups	1.1	.3
8. destroying property	1.2	.5
9. sex trouble	1.1	.4
10. missing dorm meeting	2.2	1.1
11. misbehaving in dorms	1.9	1.1
12. making believe you're sick	1.3	.7
13. taking other people's property	1.1	.5

Note.--N = 63.

Student Questionnaire

I'd like to talk to you about the token economy. I'd like to ask you some questions to find out how you feel about it. First I'd like to ask you some things about the school.

Student Questionnaire

Name Age Sex Tribe
Grade

1. Can you remember and describe how you felt about coming to this school?

2. What do you think of the school in general?

3. What are the good things about being a student here?

4. What are the bad things about being a student here?

5. What do you think of your teachers?

6. Which teacher do you like best?
   a. What do you like about him or her?
   b. What don't you like about him or her?
7. Which teacher do you like the least?
   a. What don't you like about him or her?
   b. What do you like about him or her?
8. What are the people who work in the dorm like?

9. Which dorm person do you like best?
   a. What kinds of things does that person do which you like?
   b. Does he (she) do anything you don't like? What?
10. Which dorm person is the worst?
   a. What kinds of things does he(she) do that you don't like?
   b. Does he(she) do anything you do like? What?

11. How did you first find out about the token program?

12. Who told you about it and when?

13. What did they tell you?

14. Did anyone ask for your opinion on the program?

15. Who are the people who give you tokens in the school? -- What for?

	Names	Behaviors
1.		
2.		
3.		
4.		
5.		
6.		
7.		
8.		
9.		
10.		

16. Who are the people who give you tokens in the dorm? And what for?

	Names	Behaviors
1.		
2.		
3.		
4.		
5.		
6.		
7.		
8.		
9.		
10.		

Of these, which are the most important?
Of these, which are the least important?

17. Who gives you the most tokens?

Why?

18. Who gives you the least tokens?

Why?

19. Do you always get tokens when you think you should get them?

    If not, how come?

20. Do you think the token program is good for this school?

    Why? or why not?

21. What kinds of things have you gotten with your tokens?

22. Have you gotten anything you really wanted?

23. If so, did you save for it?

    How long?

    How many tokens?

    If you haven't gotten things you wanted, why haven't you?

24. Is there any privilege or thing you want a lot but you can't get with tokens?

    Example: food
    trips
    T.V.
    a recreation activity

25. Do you feel the teachers or dorm people try to control you using the token economy?

    Give an example

26. Whose idea was it to start this program?

27. Why did they start it?

28. How did they pick the behavior to earn tokens?

29. What behaviors did they pick?

    a. Of these, which are the most important?

    b. Of these, which are the least important?

30. How do you think they picked how many tokens for each behavior?

31. Who decided what you were able to get with the tokens?

32. Do you think this is a good program for this school?

33. What would happen if you came in drunk to the dormitory?

34. What would happen if you disobeyed a teacher's request to be quiet?

Lets pretend you are in charge of a new token program

35. Who would you have help you plan your program?

36. How would you decide on what behavior to try to change with tokens?

37. What behaviors do you think you'd pick? Why?

38. Would you use tokens or something else?

39. If you were using tokens, what kinds of things or privileges do you think the tokens could purchase?

40. How would you get the students to understand the program before it began?

41. These are the behaviors the program was trying to change. Rate yourself on them before and after the program.

	Before					After				
	Never	Seldom	Some times	Often	All the Time	Never	Seldom	Some times	Often	All the Time
1. missing classes	1	2	3	4	5	1	2	3	4	5
2. lateness to class	1	2	3	4	5	1	2	3	4	5
3. misbehaving in class	1	2	3	4	5	1	2	3	4	5
4. doing well in school	1	2	3	4	5	1	2	3	4	5
5. drinking alcohol	1	2	3	4	5	1	2	3	4	5
6. leaving grounds without permission	1	2	3	4	5	1	2	3	4	5
7. missing bedcheck	1	2	3	4	5	1	2	3	4	5
8. picked up by police	1	2	3	4	5	1	2	3	4	5
9. destroying property	1	2	3	4	5	1	2	3	4	5
10. getting in trouble because of sex	1	2	3	4	5	1	2	3	4	5
11. missing dorm meeting	1	2	3	4	5	1	2	3	4	5
12. misbehaving in dorms or on trips	1	2	3	4	5	1	2	3	4	5
13. making believe your're sick	1	2	3	4	5	1	2	3	4	5
14. taking other peoples' property	1	2	3	4	5	1	2	3	4	5

Section V

Assessment of Staff Attitudes and Perceptions

by

Wayne P. Moellenberg, Ph.D.

Assessment of Staff Attitudes and Perceptions

The following section will be devoted to a consideration of possible relationships between conceptual systems theory, as posited by Harvey, Hunt and Schroeder (1961), and some of the findings of the evaluation team using the Teacher Interviews as a source of data.

The most striking finding that emerges when one examines the data gathered with the "This I Believe" instrument, which is an open-ended instrument devised by O. J. Harvey and his colleagues (Harvey, Hunt and Schroeder, 1961) to ascertain the conceptual system orientations of respondents, is the overwhelming preponderance of System 1 among the staff at the school. This finding came as no surprise, since numerous studies involving conceptual system orientations among educators at various levels had found System 1 orientations much more prevalent than other systems (Moellenberg and Williams, 1969; LeMar, 1968; Wangler, 1969; Moellenberg, 1971). However, none of these studies found the extent of System 1 dominance uncovered in this study, in that proportions ranged from 40 to 60% System 1 in the previous studies, while in this instance only four representatives of other systems were found among the 50 staff members who gave usable responses (92%). Among the teaching staff, there were 10 instances of pure System 1 orientation, one admixture of Systems 1 and 4, one admixture of Systems 3 and 1, one clear example of System 3 and no representatives of System 2 or System 4. An even more pronounced predominance of System 1 was found among the dormitory staff, where there were 22 clear examples of System 1, three admixtures of Systems 1 and 3, two admixtures of Systems 1 and 4, one admixture of sys-

tems 2 and 3, one clear example of System 3, and no instances of System 4.

The reader who is unfamiliar with the tenets of conceptual systems theory may wish to refer to the 1961 book (Harvey, Hunt and Schroeder, op. cit.) or some of the numerous studies conducted by these men and their students. Clear descriptions of the characteristics posited for the four systems can be found in these sources, and a good understanding of these characteristics would be of value in the discussions that follow. Very briefly, however, System 1 is defined as one involving concreteness, authoritarianism, dogmatism, rigidity, and a conservative attitude with regard to change. System 2 is characterized by rebellion against the authority upon which System 1 individuals rely, together with a strong desire for change and slightly less concreteness than System 1. System 3 falls toward the abstract end of the concrete-abstract continuum, and individuals functioning with this system are oriented toward close interpersonal relationships. They wish to avoid conflict, and are quite willing to follow the directions of one in authority. They also are much less rigid and dogmatic than individuals in System 1 or System 2. Finally, System 4 individuals are the most abstract of the representatives of any of the four systems. They are able to analyze complex issues for themselves and arrive at decisions without depending upon or rebelling against the directions of those in authority. They are flexible and open to change, which enables them to admit new evidence and modify their behaviors accordingly.

Keeping in mind these characteristics, and others not revealed in these very brief descriptions, certain observations of the evaluation team can be explained very readily. Specifically, the suspicion, resentment, and even hostility felt by those conducting interviews could have been predicted on the basis of the particular combination of circumstances and staff charac-

teristics that exist at the school.

Staff members at AIS face an exceedingly difficult task. They bear the responsibility, 24 hours every day, for controlling the behavior of several hundred rebellious young students who are made especially difficult to manage because of the tremendous adjustments that must be made in coming from the reservation to a residential school in a large city. To complicate the task still more, the past two or three years have seen an increasing percentage of these young people who are sent to the school because of family difficulties, academic problems, social maladjustments, and other similar handicaps. In short, controlling and educating these young people would be an exceedingly difficult task for anyone under the very best of circumstances.

Unfortunately, the circumstances during the past few years have been far from ideal. The staff has received very little special preparation for dealing with large numbers of disturbed youngsters experiencing culture shock. Further, numerous studies and experiments have been conducted in the effort to find new ways of coping with these problems, and these attempts have been perceived as adding to the burdens of the staff without providing the kind of clear-cut solutions they desire. In fact, some members of the staff, perceiving the changes that have appeared in the students and not fully understanding the causes, lay part of the blame on the theories and experiments that have been tried.

It certainly is not surprising, then, that a staff composed primarily of System One individuals would show some resentment toward yet another group of researchers evaluating an experimental effort. Being somewhat predisposed toward resistance to change by basic orientation, and feeling the threat of rapid change not only in the nature of their task but also in the method of accomplishing that task, it is quite understandable that they would react negatively.

Conceptual systems and semantic differential. In analyzing the responses made by the staff, it was a matter of some interest to determine whether there was a relationship between conceptual system orientation and other variables. In particular, possible relationships between system orientation and attitudes toward the referents presented in the Semantic Differential were checked.

One such relationship that might be suspected would be a tendency for System 1 and System 2 respondents to reveal stronger feelings about such matters as "Praise," "Punishment," "Adult Control," "Token Economy," and other referents presented in the Semantic Differential. In view of their presumed tendency toward dogmatism and strongly-held beliefs, System 1 and System 2 respondents might be expected to choose extremes on the scales more frequently than representatives of Systems 3 and 4.

In view of the overwhelming preponderance of System 1 respondents and the lack of representatives of other systems, it was necessary to modify the testing of the hypothesis. Therefore, clear examples of System 1 and admixtures involving System 2 were treated as one group, while admixtures of Systems 1 and 3 or Systems 1 and 4 were treated as the other group. A chi-square analysis was conducted to ascertain whether there was a significant difference between these two groups in terms of the number of strong positions chosen on the Semantic Differential scales. Results are presented in Table 1.

The results shown in the above table seem to indicate that conceptual system orientation is indeed related to the strength of beliefs held about the referent words in the Semantic Differential, which in turn are relevant to the Token Economy experiment.

Semantic differential. In order to ascertain the feelings of members of the AIS staff toward the Token Economy experiment, a Semantic Differential (Appendix C) containing the following referents was employed: (1) Adult Control; (2) Material Reward; (3) Praise; (4) Punishment; (5) Discipline; (6) Token Economy; (7) Authority. Respondents were asked to rate each of these referents on eight scales designed to measure the Evaluative dimension. Additionally, three referents, "Token Economy," "Praise," and "Punishment" were presented a second time with four scales to measure the Potency dimension.

In order to ascertain the relative popularity of the various referents, numbers from 1 through 7 were assigned to the responses, with high numbers assigned to favorable responses and low numbers to unfavorable responses. By this method, a total score could be computed for each individual on each referent. Also, it was possible to compute the mean rating for each referent, so that the general feeling of the group about each referent could be ascertained.

In descending order, the staff ranked the referents as follows:

1. Praise ($\bar{X}$ = 6.02)
2. Material Reward ($\bar{X}$ = 5.48)
3. Adult Control ($\bar{X}$ = 5.22)
4. Discipline ($\bar{X}$ = 5.08)
5. Authority ($\bar{X}$ = 4.96)
6. Token Economy ($\bar{X}$ = 4.65)
7. Punishment ($\bar{X}$ =3.92)

Among the second group, the ranking was the following:

1. Praise ($\bar{X}$ = 4.70)

2. Punishment ($\bar{X}$ = 3.80)

3. Token Economy ($\bar{X}$ = 3.78)

It should be noted at this point that such a procedure provides only a rough indicator of the interaction of two complex factors, in that the two dimensions tapped by the Semantic Differential are quite distinct. A referent like "Punishment," for example, ranks higher on the Potency dimension than on the Evaluative dimension, while the order is inverted for a referent like "Praise." However, since the mean takes all dimensions into account, and since both administrations produced very comparable rankings of key referents using different scales, it seems reasonable to assume that a ranking by mean score has some validity as a way of reflecting the general feelings of respondents about the relative merits of the referents presented.

Additional complexity is added to the problems in interpretation by the additional factors of social desirability and response tendency. Some individuals may have felt constrained to respond in certain ways because of presumed associations between the research team and the AIS administration, in spite of repeated assurances to the contrary by the evaluation team. Some may have been influenced in their responses by the general sense of frustration, noted earlier, that seems to have developed because of the increasing difficulty of their task. Some individuals clearly tended to choose extreme responses on the scales, while others tended toward more neutral responses. Once again, however, the use of means and a relative ranking procedure helps to provide some basis for discussion, even though it would be dangerous to assume complete validity of any scores as absolute values.

With all of these cautions, it seems safe to say that the Token Economy system has enjoyed only limited acceptance by the AIS staff. To be sure, the mean response to the term is on the positive side of neutral, and it seems quite possible that the majority of feelings on the part of the staff are

actually more favorable than unfavorable. However, the fact that the term "Token Economy," ranked sixth among the seven presented in the first round, and third among the three in the second round, indicates something much less than enthusiastic reception.

There appear to be many reasons for the lack of enthusiastic support for this program, some of which have already received mention in preceding discussions. A sense of discouragement at the enormity of the task, resentment toward another new idea that fails to provide clear-cut solutions, doubts about the rightness of a system that some regard as "bribery," lack of special training and resultant inability to use the system effectively, and many other factors as yet unknown are undoubtedly at work. Whatever the cause, however, it is clear that many of the staff are not in a position to provide positive support for the program.

On the other hand, one must not fall into the error of thinking that there is any overwhelming feeling among the staff that the idea should be abandoned. When asked how they felt about the program at the beginning, and in a second question, how they feel about it now, the staff was almost evenly split in their feelings. These data, broken down into teaching staff and dormitory staff, are presented in Appendix C.

Examination of these results indicates some trends in the data that reflect on the discussion at hand. They seem to indicate that teachers were more extreme in their initial feelings about the Token Economy than were members of the dormitory staff, but that shifts from neutral to negative positions by six members of the dorm staff make the two groups very comparable in present opinions. Also, the data indicate that younger members of both staffs tended to be more positive in their initial reactions to the project than were persons 51 or older, and that the younger group also showed more

positive shifts (or fewer negative shifts) in opinion than did older respondents.

Females were significantly more negative than males in their initial reactions, with no significant differences between the two groups in terms of changes in feeling after experience with the system.

Staff members with five or fewer years of service with AIS were significantly more likely to have positive feelings about the project than were those with six or more years and also tended to remain more positive.

Staff members with a college degree were more positive about the project than those with less education, while there were no significant differences between those two groups in the opinion changes that took place after exposure to the project.

These findings should be kept in mind through all of the following discussions, since the patterns they show may help to explain staff attitudes toward various aspects of the Token Economy program.

Teacher Interview--Part I

Item 1 on the Teacher Interview Questionnaire was the direct question, "What is the role of this school?" As might be expected, the most popular single response to that question was the equally direct answer, "Education." Twenty of the 50 persons who responded to the question responded in that way. The remaining 30 individuals either elaborated and presented additional views of what the school should do, or denied that education was the real purpose. Eleven of the 30 individuals, constituting the second largest single group, expanded the education idea to include the role of socialization agent. Four others said that it was for the purpose of education and vocational training, while seven said it was a combination of education and the provision of special help for the student. Four respondents were con-

cerned about the idea of teaching values and pride in the Indian heritage of the students, and the last group of four individuals took the rather pessimistic view that the school had no real goals and that the many problems and "problem kids" precluded an effective educational process.

The pattern of responses found with this item includes no major surprises. It appears to parallel the pattern one might have expected from any group of experienced teachers in almost any school with the possible exception of the eight individuals who talked about aspects other than the educational function. One might be somewhat surprised at the large number of respondents who viewed education as the single role of the school, but perhaps this can be explained on the basis of their possible perception that the question asked for the primary role.

Item 2 contained the request that the respondent describe the student body in reference to a number of areas, the first of which was the background from which the students came. There was considerable diversity among the respondents in terms of the kind of background information that they considered most relevant. Eleven of the staff members focused their responses on a discussion of the tribal area from which their students came, while 21 of them were concerned with what they classified as "the poor home situation" of many students. It would have been interesting to know whether this evaluation was related to a specific cultural bias, or whether it is really an objective analysis of the teachers relating to conditions that everyone could agree would be detrimental to the students. The evaluation of the responses provides little evidence about this question, with the exception of references to poverty and lack of various material comforts. Also, some teachers explained that the home situations were not conducive to academic achievement on the part of their students, which would provide some additional insight

into the basis for their evaluation.

Six individuals were concerned with the idea that many of the students were "problem children" who were sent to the school for management and therapy purposes, while three additional respondents indicated that the students were generally academically poor. Eight others felt that many students came from a "poor social background," but three others said that the students were not extremely underprivileged. Finally, two of the staff members perceived the pupils as "helpful, friendly, and eager." Two of those questioned said they did not know about the background of the students.

It is interesting that such a wide spectrum of perceptions could be gained about the backgrounds of the single student body. Of course, part of the difference would undoubtedly lie in the fact that different staff members have contact with different children, but it also seems evident other factors were at work influencing the perceptions. Granted that lack of economic prosperity, serious academic problems, and difficulty in social adjustment characterized many of the children, it is still interesting to find that home situations or social backgrounds would be characterized as "poor." At the very least, such value judgments would ordinarily be qualified by further explanations of the cultural standpoint from which the evaluation was made, especially in view of the enormous cultural transitions with which the youths are faced. It would appear that the majority of the staff are not particularly sensitive about the debate that has surrounded the whole issue of assimilation into the majority culture versus maintenance of the traditional cultural patterns.

Last, but not least, one should note that some individuals viewed these students as "problem children," while others felt that the students could be characterized as "helpful, friendly, and eager." Again, it is impossible to

tell whether this was due primarily to their exposure to different groups of children or to individual differences in the way they perceived the same children, but it indicates beautifully the wide range of viewpoints that develop even at a single resident school.

In answer to a question about the trends of the background skills of these students, by far the most frequently mentioned were such matters as their art work crafts, and handwork. Next on the list were their Home Economics, industrial arts and mechanical skills, which were mentioned by six people in comparison with 31 persons who noted the first category. Three persons noted again in this context that students were academically poor, while three others mentioned their sports and recreation, together with their ability to deal with animals. Finally, nine respondents felt that the students had no skills or very weak skills, and that they did not use their potential abilities.

These responses reflect once again the extreme difficulty encountered in separating objective realities from preconceived notions. Certainly, many students would be skilled in the art works, crafts, and handwork through which many tribes supplement their livelihood. However, in view of the influence of stereotypes, one might wonder whether such characteristics would be sufficiently outstanding to create such an overwhelming impression on the staff. It seems much more likely that this response indicates a definite interaction between the notions of the observers, influenced by stereotypes, and the objective characteristics exhibited by the students.

Defects mentioned by the staff in the backgrounds of the students ranged from difficulties in the communication arts to such matters as social maladjustment, physical defects, discipline problems, poor backgrounds, unusual behavior, lack of motivation, and lack of opportunity. The most frequently

noted were the difficulties in the communication arts and the bad discipline problems. Three staff members, on the other hand, said they did not know of any defects, or that the students had none.

The pattern found here supports earlier observations that staff members tend to note things that make their particular tasks more difficult. The difficulties they encountered in teaching the students because of problems in the communication arts, together with the very natural tendency to focus on bad discipline problems and lack of motivation, would surely support that kind of view. Little more needs to be said about this situation, except to note that there is objective support for the views of the staff in these patterns.

The interests of the students were viewed by the staff as falling primarily in the areas of arts and crafts, social life, entertainment, vocational work, and school work. The most popular choice among these categories was the entertainment area, including sports, recreation, music, television, etc., which was chosen 24 times as compared with the next largest category, arts and crafts, which was chosen 10 times. Nine individuals said that the students had no interests, or at least that they were not developed, and a very few indicated that their interests were "instinctual" or that the students were interested in being "being Americanized." Only six chose the category of vocational work and school as important interests of the students.

Once again, the pattern is not especially surprising in terms of the entertainment, social life, or arts and crafts categories, since these might be supported by both preconceived notions and objective realities. It is somewhat surprising that nine staff members would see the students as having no interests, since the normal enthusiasms of youth would usually reveal at least transient interest in a variety of things. The "instinctual"

interests category included such matters as sex, food, and other similar items, so undoubtedly some students display interest in these areas.

When asked about the goals of their students, the staff tended to be rather pessimistic. Seventeen of them said that the students intended to finish high school, and that some would go higher, but 18 indicated that the students have few goals or that their goals are so vague as to be nonfunctional. Eight persons felt that their students had a goal of a career or trade, while 11 thought student goals were in the areas of making a living, a home, and a marriage. Only two thought the goal of the students was to go back home, while one thought they wanted to "get something free from the government."

The total impact gained from the responses to Item 2 of the questionnaire is somewhat bleak. It is evident that a great many members of the staff view their students as disinterested, unmotivated, unskilled, and generally lacking in clear-cut goals. Once again, one must note that there is objective evidence to support some of these contentions, but the real crux of the question lies in the areas of cause and effect. One must ask whether such characteristics are truly embedded in the students when they arrive at the school, or whether the expectations of the staff members help to instill such characteristics. There is no question that circumstances quite outside the control of the staff members contribute greatly to the development of undesirable characteristics among the students that would militate against easy adjustment of the students to the school. The only question is whether more favorable attitudes on the part of staff members would help to lessen the impact of these negative influences. It would appear that some effort might be directed toward finding out whether this could be the case.

Item 3 approached the same sorts of questions as Item 2, except from a slightly different point of view. Staff members were asked to list the learning

assets, learning deficits, or other learning characteristics of the group of students. On the question of learning assets, the most frequent response was again the arts and crafts skills that staff members thought they noted among the students. Twenty-one persons answered in that vein. However, 14 persons could not answer or did not know of any special learning assets. Eleven others listed motivation, imagination, and capability as learning assets that could be found among the student body, while the others mentioned bilingualism of their students and the rich cultural backgrounds from which they came. A few others listed such things as sports or social graces, together with such things as academic skills, as important contributors, but this question seems to be difficult for a number of staff members.

In the area of learning deficits, there was considerably more agreement. Twenty-eight persons specified academic deficits, especially in areas of English and other communication arts. Lack of motivation and attention was specified by 14 other persons, while background deficits were listed by nine more. Eight said they did not know of any specific learning deficits and two said that the morale of the staff was a problem.

When asked to list other problems not specified by the interviewers, most of the respondents could not venture specific suggestions, with the result that 37 gave no response. Among those who did respond, two or three votes were given to matters like lack of parental cooperation, too many extra-curricula activities, and the general deterioration in student behavior that some staff members thought they had noted in recent years.

Conclusions about the responses to this item would be very similar to those drawn on the basis of the preceding item. Considerable discouragement was revealed in the responses of some staff members, but it was not possible to ascertain whether this is due primarily to objective reality or to perceptions.

Item 4 attempted to tap directly into teacher perception of their problems at the Albuquerque Indian School. In response to the direct question, 16 teachers indicated the childrens' behavioral patterns were the source of their most difficult problems. Six others said they had no problems, while eight additional persons were troubled by lack of communication. Another eight persons said that the diversity of background among their students was troublesome and ten said they had problems with the administration. One person had social problems in the school, seven others gave no response at all to the item, three said they were troubled by trying to hlep the students learn, and two said there was not enough discipline.

The pattern formed by these responses would certainly conform to expectations. Once again, there is evidence that behavioral problems and discipline loom large in the eyes of staff members. As in any institution, this results in a certain amount of friction with the administration, since staff members are likely to expect more help with such problems than they are able to obtain. The lack of communication that has been mentioned previously in connection with student deficits arose again as a source of teacher distress. About the only real surprise, in the light of all the previous discussion of problems and trials, would be the fact that six persons said they had no special problem as teachers at the Albuquerque Indian School.

Closely paralleling Item 4 was Item 5, which asked for some indication of the unique rewards for a teacher at the AIS. In response to this question, 33 persons indicated that personal satisfaction achieved through teaching and helping were the unique rewards they had experienced. Nine persons indicated, rather dejectedly, that there were no rewards obtainable from working in the school, while three others said rewards were limited to the financial area. Two persons were rewarded by the sense of helping their own tribe

and five others were unable to give a response to that item.

Since the socially desirable response is rather evident here, it is difficult to ascertain the extent to which this item tapped the true feelings of the staff. However, in spite of the frustrations and difficulties, it is clear that enormous opportunities for personal satisfaction should be available at the Albuquerque Indian School. Therefore, it seems only reasonable to take the responses at face value and assume that the staff gains great satisfaction from what they are able to accomplish in spite of the difficulties. Probably the most troublesome aspect of the situation is the fact that nine individuals felt there were no real rewards in their work. While still a distinct minority, that group is sufficiently large to call for some careful attention in terms of staff morale. Responses to this item would then clearly support previous conclusions that additional work with the staff is needed, possible including some additional incentives along with additional specific training.

Item 6 explored the feelings of the staff in terms of perceived distance from others with whom they would work. First, staff members were asked to indicate whether they felt closest, in terms of interests and goals, to colleagues, administrators, consultants, students, or parents. Then, with regard to the same groups, they were asked to indicate from whom they felt furthest.

Results from the first question indicated a near tie between colleagues and students. Twenty-two individuals felt closest to their colleagues, while twenty-one felt closest to their students. Only five felt closest to the parents, while three felt closest to the administrators and one felt closest to the consultants.

The group from whom the staff felt farthest removed were the administra-

tors, who were chosen for that role by twenty respondents. Parents came second with 16 votes, while consultants were third with ten, colleagues were fourth with four and only one individual felt far removed from students.

These results are interesting, but disturbing. There is no surprise or distress in the fact that the staff members felt closest to their colleagues and students. That would be a natural and healthy situation in any school, if for no other reason than the amount and kind of contact that occurs. However, it is not encouraging when staff members indicate that they feel farther from administrators and parents than from outside consultants. In the light of other evidence already evaluated, this indicates beyond any question the serious need for extensive public relations work to build better communication between administrators and staff members. It also provides clear evidence that work is needed to build closer relationships between parents and staff members. Such work has already been attempted, but apparently with no great success. Perhaps additional funds should be directed toward these areas in the attempt to build better relationships.

Teacher Interview--Part II

Part II of the Teacher Interview was designed to get at specific aspects of the Token Economy Project, in contrast to Part I which dealt with background information about the entire school.

Item 1, which related directly to the Token Economy Program, asked about the manner in which staff members first learned about the project. Staff members were asked who informed them, along with when and where they were told.

It was evident from the responses that some internal problems may have been exacerbated by lack of opportunity for adequate public relations work to prepare the staff for the project. The largest single group (25) first learned about the project at the workshops that were held to provide infor-

mation and training for the project. There were 13 others who said they heard about it from Mr. Blanchard, and the remaining 12 persons named a wide variety of sources ranging from newspapers to colleagues. It was clear that some who obtained their information indirectly seemed rather bitter about that fact.

The time when staff members learned what was going on also seemed quite variable. Most (37) found out in the Summer of 1970, but a few had some knowledge already in the Spring of that year while others were not aware until Fall.

In view of the variability in mode and time, it was not surprising that the place also varied. The largest single group (21) identified AIS as the place where they heard the news, but 20 others placed the location at the Albuquerque Academy. One was told in Santa Fe, two in Flagstaff, one by newspaper, and five could not recall.

The kind of generalization that would seem supportable on the basis of data from this item would be that greater uniformity in time and mode of informing staff would have been better in terms of staff morale. This should not be regarded as a criticism of the project directors, since the problem was caused in large part by the fact that final approval of specific plans was delayed. However, it is clear that the second year of the project could benefit greatly from greater uniformity and systemization of the process through which staff members are apprised of future plans. Needless to say, staff involvement in the planning process, when appropriate, would also be of great benefit in this regard.

Item 2 dealt directly with the point just made, in that interviewees were asked directly what part they had in setting up the program. As expected from previous responses, the majority (33) said they had no part.

Three others said they decided what behavior to reward and when, four said they set up their own system contracts, three said they helped to plan goals, four were consulted about the program, one simply agreed, and two helped to decide the worth of tokens.

Again, background information is essential in order to evaluate the meaning of such responses. Obviously, the timing problems already discussed also precluded the involvement of many staff members, since they were not available for consultation when final planning was accomplished. Also, the specialized knowledge necessary for the planning task and the tight time schedules militated against committee-type planning in this project. However, since time and basic familiarity are now greater than last year, it seems that some form of task-force orientation for dealing with special problems might help to increase the sense of involvement that is so vital to the success of the project.

The training given for the program was the focus of Item 3. The evaluation team was interested in learning who had been involved in the process and how staff members perceived their training.

It was clear from the responses obtained that uniformity had not been achieved in this phase of the operation either, since 18 respondents said that no one had trained them, 13 identified Mr. Blanchard, one mentioned Mr. Adams, seven said they had been trained by SWCEL, three by their supervisors, one by Mr. Reedy, two said they did not remember, and eight gave no response. A caution is again in order, in that some of the diversity is a function of differences in staff assignments and backgrounds. However, the fact that such a large number would indicate that they had received no special training for this project is troubling. Very few had any previous training or experience in the principles of systematic behavior modification, so a lack

of project training would mean that they were left to their own devices in attempting to implement the Token Economy.

Evaluations by staff members of their training reflected this situation, in that 31 of them said their training was bad or non-existent. On the other hand, 12 found the training good or very good, and seven failed to respond. Such a pattern clearly supports the suppositions of the preceding paragraph, although it should be noted that 12 individuals were decided exceptions to the general rule.

When asked if they could tell why the training was inadequate, most staff members criticized amount rather than quality. Fourteen repeated again that they had received none, while 18 others said it had been too short. Only one blamed a negative attitude on the part of the person providing training, while seven thought it was disorganized and confused. Two blamed lack of competence among trainees for the problem, and two others felt that lack of consultation with the staff was at fault.

The total picture that emerges from Item 3 is quite clear, and again supports many of the earlier conclusions. Lack of systemization in the training phase, undoubtedly related to the same lack of time for planning and preparation mentioned previously, seemed to exert negative influence on the perceptions of staff members, and very probably also reduced their effectiveness in behavior modification. A well-organized training phase would certainly seem appropriate at this time in order to maximize the effectiveness of the staff for next year.

Data from Item 4 which asked how the respondent felt about the program originally, and Item 5 which asked how the respondent felt about the program now, were discussed in the preceding section. Referring to the data in Appendix C, Tables 2 and 3, it is obvious that great differences of opinion existed

among the staff in terms of their feelings toward the Token Economy program. Summarizing the results one can see that 23 members of the staff had positive feelings, 11 were neutral, and 15 felt negative at the beginning of the experiment. At present, on the other hand, 20 feel positive, eight are neutral, and 21 feel negative.

With such diversity of feeling, along with other problems faced by the project it is not surprising that responses to almost all items related to staff perceptions show lack of concensus.

The obvious implication of such findings, of course, is that intensive public relations work with the staff will be necessary in order to generate the kind of emotional support required.

Since the evaluation team had some indications that many staff members were not consistently using the token system, Item 6 was included in the interview to ascertain whether that might be the case. It proved to be so, since only 18 persons indicated that they employed the system "all" the time, 16 said they did "most" of the time, ten "some" of the time, two "very little" one "none," and three said the question did not apply to their situations. Clearly, this constitutes anunfortunate situation, in that consistency among the staff would seem essential to the successful functioning of the enterprise. Lack of consistency would seem likely to cause frustration and resentment among students, since some would be rewarded for given behaviors and others would not receive reinforcement for the same behavior.

The implications of this finding are that staff members need to be convinced of the merits of the program and the necessity of consistency in application. This, in turn, would support the need for public relations work with the staff, additional training, the establishment of clear guidelines for utilization, and possibly an incentive program for staff members.

Item 7 dealt with much the same area of concern as Item 6, except that it asked specifically for average rate of reinforcement instead of the general utilization of the program which was the concern of Item 6. The same variations were found, in that 23 persons indicated they gave tokens "once per week," 11 said "once per day," two "every period," five "at the time the desired behavior occurs," three said the question did not apply to them, and one person did not give tokens. Such responses indicate that not only are there differences in the extent to which individuals employ the system, but also in the way they use rewards. Some staff members seem to be on fixed-schedule systems which differ in time among individuals, others seem to be on fixed-ratio (piecework) systems, and still others seem to be using variable-schedule or variable-ratio schedules for reinforcement. If so, confusion and resentment among students would be expected, since disappointments and frustrations would be numerous.

The same implications (need for "selling," training, and rewarding staff) would seem to follow from these data as from those in the preceding item.

The diversity among staff members was nowhere more clearly reflected than in Item 8, which asked about the kinds of behaviors the program was successful in shaping. Predictably some said "none" and others said "all kinds." Additional exploration reveals, however, that many staff members had noted some differences in their relative success in changing different behaviors. Some said, for example, that tardiness and absenteeism seemed to be modifiable with reinforcement techniques. Four classified "desirable" behaviors as easier to shape than "aversive," probably because of the fact that they can be directly rewarded.

In answer to Item 9, which asked about the kinds of behavior that the Token Economy Program is not effective in changing, nine somewhat disgruntled

staff members responded with "every kind." Some of them went on to state their opinion that the program was a waste of money and that their efforts could be more productively spent in other kinds of endeavor. Twenty other individuals said the program was not good for controlling undesirable behaviors and discipline problems, while 16 said it was not good for controlling desired behavior and achievements. One staff member asserted that such a program goes against the Indian way, one said that there were no behaviors that he had found which could not be influenced by such a program, one said that cognitive behaviors were not responsive to the Token Economy System, two said that value judgments were not influenced, and five said that they did not know how to answer the question.

The fact that so many teachers identified undesirable behavior, discipline problems, and achievement as areas not responsive to management with the Token Economy System is interesting. The evaluation team had arrived at this conclusion prior to the analysis of the item, since deep-seated and complex syndromes are difficult to treat and hard to manage in a reward ysstem. Staff members had evidently arrived at the same conclusion on the basis of their experience, and had analyzed the situation to the extent that they understood many of the reasons for the difficulty.

Item 10 was another direct question asking staff members to indicate how well they thought the program was working at Albuquerque Indian School. In response to this question, seven persons said that it was not working at all, while 19 more felt that it was working very poorly. Four classified success as "fair," ten said that it was "pretty good in some areas," two said they had very good success, three said things were going backward, and five people indicated no opinion.

While results such as these are distressing, in the sense that they

indicate a sizeable proportion of very negative perceptions of the program on the part of the staff, conditions may not be as bad as they look. In the first place, it became evident through responses to several questions that a few staff members were simply responding negatively to all aspects of the Token Economy Program, without sorting out the various aspects. Therefore, while these responses should not be discounted, they can be treated as part of a more generalized resistance to the program rather than specific evidence of failure. With regard to those who said that things were going very poorly, or even the three who said they were going backward, it seems evident that another factor was at work. The comments that accompany these evaluations indicated that many of the respondents would expect fairly dramatic evidence of change in broad areas before they would classify the program as successful. This is not an unreasonable position in view of the statement of project objectives, but it seems unlikely that such a degree of success realistically could have been expected. This may have been recognized by the individuals who said that success was "pretty good in some areas." In the same vein as Item 9, they had identified some kinds of behavior as simply too resistant to achieve change in a short period of time, while other simpler behaviors seem to show some modification.

In terms of whether the program actually helped or hindered the work of the individual staff members, 18 said that it helped and 20 said that it hindered. Four said that it did neither, two said it did both, four others said that it does not matter or makes no difference, and two gave no response. In short, it was a mixed bag.

In response to the question of why this might have been true, three identified the age of the students as a factor, indicating better success with the younger children than with older. Fifteen said that the attitudes

and understandings of the staff were poor, resulting in inconsistent management of the program. Two said the trouble was due to conflict with the value system of the Indian culture, while seven said that the system simply does not work, presumably meaning that behaviors cannot be changed with token rewards. Several others mentioned such things as lack of supplies, lack of guidelines, student misunderstanding or resistance, and similar problems.

Item 11 asked all staff members "How much of the time do you give praise along with the tokens?" Responses to this question revealed some interesting observations with regard to ratings of certain referents on the Semantic Differential.

First, it should be noted that praise was very widely given by both the teachers and dormitory staff along with the tokens. Ten of 21 teachers and five of 22 dormitory staff who responded to that question indicated that they always gave praise with the tokens. Six additional teachers and six more dormitory staff members indicated that they almost always gave praise. No teachers, and only two dormitory staff members, indicated that they rarely or never gave praise with tokens.

These observations are enlightening when considered together with the responses to the Semantic Differential referents. Most obvious, of course, is the fact that it provides additional evidence, together with the high Semantic Differential ranking given "Praise," that staff members regard praise as an effective and appropriate motivator. Indeed, since they ranked that term above "Material Reward," and far above "Token Economy" and "Punishment," it would seem that it is regarded as the most effective and appropriate modifier of behavior. Only eight individuals ranked "Material Reward" higher than "Praise," while 32 inverted the ranking and five rated them even.

These findings relate to criticisms of the Token Economy voiced by some staff members, who indicated concern about "bribing" the youngsters to obey regulations. These individuals felt that social controls should be used instead of material rewards, and that students should learn to obey regulations because it was the right thing to do. Several felt that more consistent punishment for clear infractions of the rules would be helpful, since they felt that students counted on their ability to escape the penalties prescribed by the staff.

Item 12 presented five areas or aspects of the program and asked respondents to rate whether each aspect was hard or easy to carry out. The first of these, consistency with tokens, was rated hard by two people, moderately difficult by eight others, neither hard nor difficult by 11, fairly easy by ten, and quite easy by 13. Six gave no response. If we take these figures at face value, it appears that individuals are not experiencing difficulty being consisten from one time to another in their use of tokens. It would appear that lack of consistency, which is mentioned frequently by the staff as a source of difficulty, is in terms of differences among individuals rather than differences in time.

A very similar pattern was found when respondents were asked to rate their difficulty in the use of praise. Four found it quite hard, seven moderately hard, five neither hard nor easy, nine found it quite easy, and 20 found it very easy. Five failed to respond.

A larger group of individuals experienced some difficulty identifying target behaviors. Four rated that very hard, 13 moderately hard, eight neither hard nor easy, 11 moderately easy, and nine quite easy.

One might have anticipated that structuring the situation in order to use the tokens would have been quite difficult for some of the untrained

staff members. However, only five rated that task as hard for them, while eight found it moderately hard, seven neither hard nor easy, 12 moderately easy and 13 quite easy.

Finally, three found the task of recording the behavior to be hard, ten found it moderately hard, six neither hard nor easy, 11 found it relatively easy, and 14 said it was quite easy.

Taken as a group, all of the responses to Item 12 would seem to indicate that the individual tasks involved in administering the program are not beyond the grasp of the staff members. To be sure, some found each task to be quite hard, indicating the need for some assistance and additional training. However, with the exception of noticing the target behaviors and recording the behaviors, the number of individuals who found tasks to be hard or moderately hard was always small. The majority of the staff rated such tasks as easy or moderately easy. In view of comments made in response to other items, indicating that the staff members found the program difficult to administer, it seems reasonable to conclude that the sum total of the behaviors make the task too complex, but that the individual components of the task are not especially difficult.

Staff perceptions about the feelings of the superintendent and principal toward the Token Economy Project were solicited in Item 13. Each respondent was asked to indicate whether he thought the superintendent and principal (1) strongly favored, (2) moderately favored, (3) were indifferent, (4) were moderately against, or (5) were strongly against the project.

It was found that 21 staff members thought the principal and superintendent were strongly in favor of the project. Ten though they were moderately in favor, while six felt they were indifferent. No respondents thought they were against the program, although 11 indicated that they did not know how the

administrators felt about it.

Such findings certainly do not indicate that the trouble lies in perceived lack of support for the program on the part of the administration. Since none of the staff members, including even those who strongly opposed the program themselves, perceived the administrators as opposing the program in any degree, it does not appear that reluctant staff members are drawing their cues from administrators.

To complete the picture of staff perceptions of various feelings about the children, Item 14 asked staff to evaluate how they thought the students felt about the Token Economy Project. Surprisingly, in view of earlier responses, 30 staff members thought the students liked the program. Twenty others thought the students were largely neutral and apathetic about the whole situation, while only four thought the students were against it. Two members said that some students cared and some did not, while four said that it hurt the pride of students to take tokens. Finally, four indicated that the students had no respect for the program, and that many took advantage of it.

These results, together with those in the preceding item, clearly indicate that staff members perceive themselves as the only ones who do not favor the program at least on emotional grounds.

Item 15 was inserted at this point to ascertain what the staff members did before the Token Economy Project in order to maintain the motivations of their students. Each respondent was asked to describe and evaluate the kinds of things that he did for motivational purposes.

While these responses are too complex to be summarized effectively, it is possible to say that they fell into the general categories of praise, trips, recognition, incentives, activities chosen by students, assurance,

counseling, appeals to pride, standards, grades, punishments, and enrichment of the class to make it interesting. A great many respondents felt that they had experienced good success with their own particular systems, although a considerable number indicated that their systems did not work effectively with many of the children who had recently been coming to the school, revealing once again the perceptions of the staff regarding recent trends in the student population.

For present purposes, it seems sufficient to note, on the basis of Item 15, that the majority of staff members in the Albuquerque Indian School rely rather heavily on traditional reward systems that one might find in almost any school.

When asked in Item 16, who, if anybody, might interfere with their practices in terms of the Token Economy Program, staff members were in fair agreement. Forty-four said that nobody interfered, while two said that "the kids" interfered, presumably by not cooperating with the teachers' systems. Two of the respondents indicated that a certain administrator was the principal one who might interfere with what they were doing, and another said the schedule prevented them from using the system.

The most important information to be gained from these responses is that teachers were left very much on their own to manage the token system. Considered in the light of the minimal training discussed in preceding sections, such findings might help to identify the source of considerable lack of consistency in the program. Many staff members did not perceive the administration of the Token Economy as one of their regular responsibilities during the past year, but rather perceived it as almost a voluntary task which they were free to accept or reject according to their own feelings. The result, of course, was very wide variation, and some resentment of the minimal super-

vision that was given. It seems clear that greater system recognition of the Token Economy as part of the regular school program might be necessary before staff members could be consistentwith one another, and would accept supervision as a matter of course.

In order to ascertain whether there was general agreement among staff members about the kinds of behaviors that most needed modification, each individual was asked in Item 17 to decide which five of 15 behavioral objectives (misbehaviors to be modified) should be rated as most important, which five should be rated as moderately important, and which five should be rated as least important. The results for each objective, broken down into teaching staff and dormitory staff, are presented in Tables 13 to 27 in Appendix C. In these tables, the index for each objective was computed by the very simple process of multiplying each top rating by 3, each intermediate rating by 2, each low rating by 1, and taking the sum. This process enables one to derive a total ranking of priorities assigned by the staff to the objectives of the project. That ranking is as follows:

Rank	Objective
1	Drinking
2	Academic achievement
3	AWOL
4	Attitudes
5	Stealing
6	Absenteeism
7	Vandalism
8	Inappropriate classroom behavior
9	Jail and/or D-home
10	Unruly behavior
11	Missing bedcheck
12	Tardiness
13	Missing meetings
14	Sexual behavior
15	Faking illness

Examination of the tables and the resultant ranking of priorities seems to indicate several matters of considerable importance to the project. Since

some of these may have implications for the second year of the experiment, it may be appropriate to consider them in some detail.

One of the most obvious features of the rating of priorities is the extent of diversity of feeling among the staff. Every single objective received some votes in every position, meaning that some staff would regard each as an objective of high importance while others felt that the same objective was a matter of less importance. Such differences of opinion are perfectly natural and expected, of course, among any population of human beings, unless prior interactions have produced common understandings. The point of this observation is that such group consensus appears to be absent in this instance, which leaves each individual at the mercy of his unique perceptions, when greater thrust might have been achieved with some mutually-agreed-upon priorities within the broader list of objectives.

A second observation relates closely to this matter of unique perceptions, in that some differences appear between the teacher group and the dormitory staff. This is also perfectly natural, since each group sees its own type of problem most frequently and faces the task of dealing with that type of misbehavior. For example, "Inappropriate Classroom Behavior" would be expected to loom larger in the eyes of teachers than in the view of the dormitory staff, and that expectation is born out by the data in Appendix C, Table 15. Eleven teachers saw that as a very important problem and only four saw it as unimportant. In contrast, only seven of the dormitory staff rated it very important, while 16 rated it unimportant. Such differences in perception appear fairly frequently in the list, and in the absence of clear guidelines they would operate to reduce the consistency with which the staff reacted to certain kinds of behavior on the part of students.

A third observation concerns the kind of problems identified as most

important by the staff. Again, it is only natural to expect that they would choose the most troublesome overt behaviors that make the jobs of staff members more difficult, and place less importance on behaviors in which the primary consequences relate to the individual involved. That was clearly the pattern found here, where aggressive and acting-out behaviors were rated high and the more passive behaviors such as faking illness, missing meetings, and sexual misbehavior were rated much lower. A serious problem with this approach is that such behavior as drinking, stealing, or unruliness is symptomatic of deeply-rooted problems, and if it is already an established pattern in the group it is difficult to eliminate by means of token reinforcement. The absence of the negative behavior over a period of time can be rewarded, but it is difficult to find an appropriate high-probability positive behavior to reinforce as the antithesis of such negative behaviors as drinking or stealing. It may be that smaller units of less complex positive individual behavior will need to be identified and clearly agreed upon by the staff in order to make the reinforcement concept work more effectively.

In Item 18 staff members were asked during the interviews to indicate whether they thought the program had gone right or wrong in a number of specific areas, and to tell how they thought the wrong things should have been done in order to make the program more successful. Responses were obtained in each of the following areas:

1. How it was introduced to students

The most frequent single response to this question area was that things had "gone right," which was the opinion of ten members of the staff. However, 23 individuals voiced specific criticisms, among which the most common (nine instances) was that the program had not been explained thoroughly enough. Along with this feeling, some expressed

the view that there had not been enough time, the assembly at which the idea was explained was not sufficiently structured, and that there had not been enough follow-through. There was also some feeling that students, teacher, parents, and board members should have been involved more in the planning of the project, or that their opinions about it should have been sought in advance. Also, the idea of students "being paid for things they are supposed to do" received criticism, especially with regard to the possibility that the basic concept was not thoroughly explained to the point that everyone could understand.

2. How it trained students

Responses in this area very closely paralleled those in (1) above, in that insufficiency and lack of consistency in the training of students received most frequent mention, followed by criticisms of the idea of "bribery" for things that should have been done without "pay." There was some feeling that younger students gained more than older students, partly because some older ones abused the system by stealing or bullying to get tokens. Six persons said that students had not been trained at all, while five others thought that things had gone well in that area.

3. The behaviors picked to be changed

Generally favorable responses were obtained with this item, in that 25 individuals though the choices were "right" or at least "O.K." Others thought that behaviors like attendance and punctuality should be reinforced, but that drunkenness, sexual activities and vandalism should be dealt with differently. Basic differences in value systems were mentioned, as was the idea that each teacher was in the best position to choose objectives for his or her own class.

4. The method of reinforcement

Great differences of opinion were found among staff members with this item. Twenty individuals felt the methods were right, while 13 felt they were clearly wrong and others expressed a variety of specific criticisms. Most of these criticisms revolved around the ideas of "bribery," the sense of materialism that it could develop, and the lack of transfer from such a system to life outside the school, especially in the light of possible conflict with Indian value systems. Lack of consistency in the reward system was mentioned, together with the observation that some teachers use symbolic rewards.

5. The responsibilities of the reinforcer

Nineteen individuals expressed general satisfaction with the responsibilities given reinforcers, while seven thought it was too much of a burden. Others again raised the problems of inconsistencies and differences among individuals, caused at least in part by lack of training and guidelines. Some apparently gave praise while others did not, and some reinforced groups while others focused strictly on individuals.

6. The responsibilities of the student

Responsibilities given students were judged appropriate by 14 members of the staff, while 20 expressed general or specific criticisms. Generally these related to specific abuses by irresponsible individuals who were careless, dishonest, or uncooperative, which several persons again attributed to inadequate explanation of the system. There was about an even split between those who felt that the experience helped students learn how to manage money and those who felt that it was harmful in that regard.

7. The value of the reinforcement

Generally favorable attitudes were expressed about the value of the reinforcement, with 25 persons indicating satisfaction and only two stating that the idea was basically wrong. There were criticisms of the rate of reinforcement and the items made available for sale with tokens, but these problems seemed to be less severe than the ones found in other areas.

8. The evaluation procedure

It seemed evident, on the basis of comments about the evaluation procedure, that some staff members had been less involved than others. Seventeen felt that the procedure was acceptable, while six complained about the lack of feedback, four doubted the influence of the evaluation on project planning, and three indicated complete lack of knowledge about any evaluation. Three other persons said there should be continuous evaluation and feedback to make the project work effectively.

9. Everyone doing it their own way

As might be expected, there were very mixed feelings about the degree of autonomy given staff members in terms of reinforcement procedures. There were 21 favorable reactions to that aspect, while 17 felt there should have been more standardization and orientation. Several mentioned specific problems that had arisen because of the lack of consistency that resulted from the high degree of autonomy.

In order to ascertain whether resistance to the project was specific to this particular program or more generalized in terms of new educational practices, Item 19 asked staff members how they felt about the introduction of new educational techniques to the school. Sixteen said they were in favor of the idea, if the techniques were appropriate and fair. Sixteen others were hesitant, presumably even under conditions described above. Two

said that the introduction of such ideas can be very confusing, one indicated that he would like to see more of a military structure in which these ideas could be introduced, and 14 had no comment about their feelings on that matter. Eight failed to respond.

It is evident, on the basis of this data, that new educational techniques receive much less than enthusiastic response from staff members. The extent to which this is due to the many kinds of studies and experiments that have been conducted there cannot be ascertained. Neither can it be said that such feelings are due to basic conceptual system orientation or other variables. What can be said is that the project, initiated as it was with serious problems of time and circumstances, could only be expected to receive a lukewarm reception during the first year in view of the prevailing attitude toward such innovations. It is important to remember this factor in evaluating opinions about the project that were expressed throughout this questionnaire.

Item 20, the last question in the questionnaire, asked staff members to indicate when they thought it was appropriate for the administration to ask a teacher at AIS to participate in a new program. A tremendous range of responses were obtained to this question, with no clear trends in those reponses. Twelve persons said it was appropriate any time if the new system would improve present circumstances. Four others said that it was appropriate if explained well in advance, while another group of four said that if the administration would ask instead of force them to accept the innovation, it would be all right. Five persons said it was never appropriate, while six others said it was appropriate only if teachers and/or staff were included and involved in the planning. Ten gave no response to the item, and the remainder made various kinds of tangential comments which did not directly

answer the question. These were generally comments relating to the quality, scope, timing, and other aspects of the program itself. In general, they might be summarized as a kind of "If they knew what they were doing" attitude toward innovative programs or procedures.

Summary of Conclusions and Recommendation

1. At the present time, while there is no overwhelming sentiment against the Token Economy among the AIS staff, it lacks the broad base of positive support that will be needed to make it function with optional effectiveness. There are many reasons for the lack of enthusiasm, including such diverse factors as the following:

   a. Insufficient time to involve the staff adequately in planning and setting up the program;
   b. A general sense of frustration among the staff about the increasing difficulty of their task;
   c. A feeling that the Token Economy is just another in a series of studies and experiments that provide little long-term benefit;
   d. Uncertainty about the attitudes, rules and relationships of various AIS administrators in and to the project.

2. The record-keeping system needs to be improved and simplified. The present system permits too many oversights, inaccuracies and abuses. Also, it is cumbersome for both students and teachers, to say nothing of the "black market" in tokens that threatens to develop among certain students. Perhaps some sort of credit card system can be devised to permit a careful accounting of each individual student's achievements and activities.

3. Some staff members experience uncertainty and difficulty in administering the system. It appears that additional training in principles of reinforcement, contingency management and behavior modification would be most appropriate.

4. The number and complexity of the behaviors to be modified is so great that staff members are experiencing difficulty. It may be necessary to identify a smaller number of less complex behaviors, and to agree upon small specific units of individual positive behavior that will be reinforced by all staff members.

References

Harvey, O. J., Hunt, D. E., & Schroeder, H. M. Conceptual systems and personality organization. New York: John Wiley & Sons, Inc., 1961.

LeMar, H. A study of the criteria used in teacher selection among elementary school principals. Unpublished doctoral dissertation, University of New Mexico, 1968.

Moellenberg, W. P. Characteristics of candidates for teacher education programs, in progress.

Moellenberg, W. P., & Williams, J. D. Conceptual systems orientation and the superintendent in negotiation. ISR Journal, 1969, 1 (1).

Wangler, D. G. Conceptual systems, attitudes toward, and achievement within the school. Unpublished doctoral dissertation, University of New Mexico, 1969.

Appendix C

Table 1

Number of Extreme Positions Chosen by Representatives of Differing Conceptual Systems

	Less than 30 strong positions chosen	Less than 30 strong positions chosen
System 1 or System 2	10	19
Systems 1-3 Admixture or Systems 1-4 Admixture	10	5

Note--$\chi^2 = 4.13$, $p < .05$

Table 2

Original Feelings about Token Economy, Teachers vs. Dorm Staff

	Positive	Neutral	Negative
Teachers	11	2	9
Dorm Staff	12	9	6

Note--All Staff: $\chi^2 = 8.10$, $p < .05$

Table 3

Present Feelings about Token Economy, Teachers vs. Dorm Staff

	Positive	Neutral	Negative
Teachers	9	4	9
Dorm Staff	11	4	12

Note--All Staff: $\chi^2 = 6.49$, $p < .05$

Table 4

Changes in Feelings about Token Economy, Teachers vs. Dorm Staff

	More Negative	Same	More Positive
Teachers	5	11	6
Dorm Staff	12	11	4

Note--$\chi^2 = 7.04$ $p < .05$

Table 5

Present Feelings about Token Economy, Younger vs. Older Staff

Age	Negative	Neutral	Positive
50 or younger	9	4	13
51 or older	12	4	7

Note--$\chi^2 = 8.03$ $p < .05$

Table 6

Changes in Feelings about Token Economy, Younger vs. Older Staff

Age	More Negative	Same	More Positive
50 or younger	7	13	6
51 or older	10	9	4

Note--$\chi^2 = 5.00$, N.S.

Table 7

Present Feelings about Token Economy, Males vs. Females

Sex	Negative	Same	Positive
Male	7	3	10
Female	14	5	10

Note--$\chi^2 = 7.91$, $p < .05$

Table 8

Changes in Feelings about Token Economy, Males vs. Females

Sex	More Negative	Same	More Positive
Male	7	8	5
Female	10	14	5

Note--$\chi^2 = 4.92$, N.S.

Table 9

Present Feelings about Token Economy, by Years at AIS

Years at AIS	Negative	Positive
5 years or less	2	8
6 years or more	19	11

Note--$\chi^2 = 5.65$, $p < .05$

Table 10

Changes in Feelings about Token Economy, by Years at AIS

Years at AIS	More Negative	Same	More Positive
5 years or less	3	5	2
6 years or more	13	16	8

Table 11

Present Feelings about Token Economy, by Education

Education	Negative	Neutral	Positive
Less than B.A. Degree	12	3	8
B.A. Degree or Graduate Work	9	5	12

Note--$\chi^2 = 8.13$, $p < .05$

Table 12

Changes in Feeling about Token Economy, by Education

	More Negative	Same	More Positive
Less than B.A. Degree	9	10	4
B.A. Degree or Graduate Work	8	12	6

Table 13

Objective: Absenteeism

	Very Important	Moderately Important	Not Important
Teachers	11	9	2
Dorm Staff	7	11	9

Table 14

Objective: Tardiness

	Very Important	Moderately Important	Not Important
Teachers	7	6	9
Dorm Staff	4	9	14

Table 15

Objective: Inappropriate Classroom Behavior

	Very Important	Moderately Important	Not Important
Teachers	11	7	4
Dorm Staff	7	4	16

Table 16

Objective: Academic Achievement			
	Very Important	Moderately Important	Not Important
Teachers	12	8	2
Dorm Staff	19	3	5

Table 17

Objective: Attitudes			
	Very Important	Moderately Important	Not Important
Teachers	12	6	4
Dorm Staff	15	8	4

Table 18

Objective: Drinking			
	Very Important	Moderately Important	Not Important
Teachers	13	6	3
Dorm Staff	17	8	2

Table 19

Objective: AWOL			
	Very Important	Moderately Important	Not Important
Teachers	11	6	5
Dorm Staff	14	12	1

Table 20

Objective: Missing Bedcheck			
	Very Important	Moderately Important	Not Important
Teachers	3	5	14
Dorm Staff	9	8	10

Table 21

Objective: Jail and/or D-home			
	Very Important	Moderately Important	Not Important
Teachers	6	8	8
Dorm Staff	6	12	9

Table 22

Objective: Vandalism			
	Very Important	Moderately Important	Not Important
Teachers	4	15	3
Dorm Staff	7	13	7

Table 23

Objective: Sexual Behavior			
	Very Important	Moderately Important	Not Important
Teachers	3	7	12
Dorm Staff	6	9	12

Table 24

Objective: Stealing

	Very Important	Moderately Important	Not Important
Teachers	6	13	3
Dorm Staff	10	13	4

Table 25

Objective: Missing Meetings

	Very Important	Moderately Important	Not Important
Teachers	3	3	16
Dorm Staff	3	8	11

Table 26

Objective: Unruly Behavior

	Very Important	Moderately Important	Not Important
Teachers	6	7	9
Dorm Staff	6	12	9

Table 27

Objective: Faking Illness

	Very Important	Moderately Important	Not Important
Teachers	2	4	16
Dorm Staff	1	4	22

Semantic Differential

Respondent Number _____

INSTRUCTIONS:

This is a method used to measure the meanings of certain things to various people by having them judge them against a series of descriptive scales. Please make your judgements on the basis of what these things mean to you.

Here is how you are to use these scales:

If you feel that the concept at the top of the page is very closely related to one end of the scale, you should place your check-mark as follows:

FAIR ✓ : : : : : : : UNFAIR

or

FAIR : : : : : : ✓ : UNFAIR

If you feel that the concept is quite closely related to one or the other end of the scale (but not extremely), you should place your check-mark as follows:

STRONG : ✓ : : : : : :WEAK

or

STRONG : : : : : ✓ : :WEAK

If the concept seems only slightly related to one side as opposed to the other side (but not really neutral), then you should check as follows:

ACTIVE : : ✓ : : : :PASSIVE

or

ACTIVE : : : : ✓ : :PASSIVE

The direction toward which you check, of course, depends upon which of the two ends of the scale seem most characteristic of the thing you're judging.

If you consider the concept to be neutral on the scale, both sides of the scale equally associated with the concept, or if the scale is completely irrelevant, unrelated to the concept, then you should place your check-mark in the middle space:

SAFE : : : ✓ : : :DANGEROUS

IMPORTANT:

(1) Place your check-marks in the middle of spaces, not on the boundaries:

THIS not this:

: : ✓ : : ✓ :

(2) Be sure you check every scale for every concept--do not omit any.

Note.--Semantic Differential administered to AIS staff.

ADULT CONTROL

Good_____:_____:_____:_____:_____:_____:_____:Bad

Clean_____:_____:_____:_____:_____:_____:_____:Dirty

Valuable_____:_____:_____:_____:_____:_____:_____:Worthless

Cruel_____:_____:_____:_____:_____:_____:_____:Kind

Pleasant_____:_____:_____:_____:_____:_____:_____:Unpleasant

Happy_____:_____:_____:_____:_____:_____:_____:Sad

Awful_____:_____:_____:_____:_____:_____:_____:Nice

Fair_____:_____:_____:_____:_____:_____:_____:Unfair

MATERIAL REWARD

Happy_____:_____:_____:_____:_____:_____:_____:Sad

Unfair_____:_____:_____:_____:_____:_____:_____:Fair

Good_____:_____:_____:_____:_____:_____:_____:Bad

Nice_____:_____:_____:_____:_____:_____:_____:Awful

Worthless_____:_____:_____:_____:_____:_____:_____:Valuable

Clean_____:_____:_____:_____:_____:_____:_____:Dirty

Unpleasant_____:_____:_____:_____:_____:_____:_____:Pleasant

Kind_____:_____:_____:_____:_____:_____:_____:Cruel

PRAISE

Good_____:_____:_____:_____:_____:_____:_____:Bad

Clean_____:_____:_____:_____:_____:_____:_____:Dirty

Valuable_____:_____:_____:_____:_____:_____:_____:Worthless

Cruel_____:_____:_____:_____:_____:_____:_____:Kind

Pleasant_____:_____:_____:_____:_____:_____:_____:Unpleasant

Happy_____:_____:_____:_____:_____:_____:_____:Sad

Awful_____:_____:_____:_____:_____:_____:_____:Nice

Fair_____:_____:_____:_____:_____:_____:_____:Unfair

PUNISHMENT

Happy_____:_____:_____:_____:_____:_____:_____:Sad

Unfair_____:_____:_____:_____:_____:_____:_____:Fair

Good_____:_____:_____:_____:_____:_____:_____:Bad

Nice_____:_____:_____:_____:_____:_____:_____:Awful

Worthless_____:_____:_____:_____:_____:_____:_____:Valuable

Clean_____:_____:_____:_____:_____:_____:_____:Dirty

Unpleasant_____:_____:_____:_____:_____:_____:_____:Pleasant

Kind_____:_____:_____:_____:_____:_____:_____:Cruel

DISCIPLINE

Good_____:_____:_____:_____:_____:_____:_____:Bad

Clean_____:_____:_____:_____:_____:_____:_____:Dirty

Valuable_____:_____:_____:_____:_____:_____:_____:Worthless

Cruel_____:_____:_____:_____:_____:_____:_____:Kind

Pleasant_____:_____:_____:_____:_____:_____:_____:Unpleasant

Happy_____:_____:_____:_____:_____:_____:_____:Sad

Awful_____:_____:_____:_____:_____:_____:_____:Nice

Fair_____:_____:_____:_____:_____:_____:_____:Unfair

TOKEN ECONOMY

Happy_____:_____:_____:_____:_____:_____:_____:Sad

Unfair_____:_____:_____:_____:_____:_____:_____:Fair

Good_____:_____:_____:_____:_____:_____:_____:Bad

Nice_____:_____:_____:_____:_____:_____:_____:Awful

Worthless_____:_____:_____:_____:_____:_____:_____:Valuable

Clean_____:_____:_____:_____:_____:_____:_____:Dirty

Unpleasant_____:_____:_____:_____:_____:_____:_____:Pleasant

Kind_____:_____:_____:_____:_____:_____:_____:Cruel

AUTHORITY

Good_____:_____:_____:_____:_____:_____:_____:Bad

Clean_____:_____:_____:_____:_____:_____:_____:Dirty

Valuable_____:_____:_____:_____:_____:_____:_____:Worthless

Cruel_____:_____:_____:_____:_____:_____:_____:Kind

Pleasant_____:_____:_____:_____:_____:_____:_____:Unpleasant

Happy_____:_____:_____:_____:_____:_____:_____:Sad

Awful_____:_____:_____:_____:_____:_____:_____:Nice

Fair_____:_____:_____:_____:_____:_____:_____:Unfair

TOKEN ECONOMY

Delicate_____:_____:_____:_____:_____:_____:_____:Rugged

Weak_____:_____:_____:_____:_____:_____:_____:Strong

Deep_____:_____:_____:_____:_____:_____:_____:Shallow

Heavy_____:_____:_____:_____:_____:_____:_____:Light

PRAISE

Weak_____:_____:_____:_____:_____:_____:_____:Strong

Delicate_____:_____:_____:_____:_____:_____:_____:Rugged

Deep _____:_____:_____:_____:_____:_____:_____:Shallow

Light_____:_____:_____:_____:_____:_____:_____:Heavy

PUNISHMENT

Light_____:_____:_____:_____:_____:_____:_____:Heavy

Deep_____:_____:_____:_____:_____:_____:_____:Shallow

Rugged_____:_____:_____:_____:_____:_____:_____:Delicate

Weak_____:_____:_____:_____:_____:_____:_____:Strong

Teacher Interview

Name	DOB	Sex
How long at AIS	Age	Grade and/or subject taught or Dormitory

Educational level

Where trained

How long working with Indians

Where worked previously

I. Evaluation of School

1. What is the role of this school?

2. Describe the student body in reference to the following areas:

   a. background

   b. skills

   c. defects

   d. interests

   e. goals

3. What are the special educational problems of this group?

   a. learning assets

   b. learning deficits

   c. other

4. What are the unique problems for a teacher in this school?

5. What are the unique rewards for a teacher in this school?

6. To whom do you feel the closest in interests and goals? (C)

   From whom do you feel the furthest? (F)

   ( ) colleagues
   ( ) administrators
   ( ) consultants
   ( ) students
   ( ) parents

II. Evaluate the program

1. How did you first find out about the Token Economy program (from whom when where)

2. What part did you have in setting it up?

3. Describe your training experience with the program

   Who trained you? How adequate do you think the training was? Why?

4. How did you feel about the program at first?

strongly favor	moderately favor	indifferent	moderately against	strongly against

5. How do you feel now?

strongly favor	moderately favor	indifferent	moderately against	strongly against

6. How much of the time do you use the program with the students?

all	most	some	very little	none

7. How often do you give out the tokens?

8. What kinds of behaviors is the program good for?

9. What kinds of behaviors is it not good for?

10. How well is the program working at AIS? Why?

Does the token economy help or hinder your work in the classroom or dorm? Why?

11. How much of the time do you give praise along with the tokens?

all most some very little none

12. Rate which aspects of the program are hard or easy for you to carry out

		1 hard	2	3	4	5 easy
1.	consistency with tokens	1	2	3	4	5
2.	praise	1	2	3	4	5
3.	noticing all target behaviors	1	2	3	4	5
4.	structuring the class or dorm for the tokens	1	2	3	4	5
5.	recording the behaviors	1	2	3	4	5

13. How do you think the superintendent and principal feel about this program?

strongly favor moderately favor indifferent moderately against strongly against

14. How do you think most of the students feel about this program?

15. What systems other than the token have you used to motivate students in the past? Describe and evaluate.

16. Who, if anybody, interferes with what you're doing in terms of the token economy? How?

17. This is a list of behaviors the token system tries to influence. Mark whether you think the token system has

I improved
S kept the same
W made worse

the behavior.

( ) Absenteeism
( ) Tardiness
( ) Inappropriate classroom behavior
( ) Academic achievement
( ) Attitudes, self concept and coping behavior
( ) Drinking
( ) AWOL
( ) Missing bedcheck
( ) Jail and/or D-Home
( ) Vandalism
( ) Sexual behavior
( ) Stealing
( ) Missing scheduled meetings
( ) Unruly behavior in dorms and on trips
( ) Faking illness

18. Next I'd like you to look at these objectives again and sort them into three groups, very important, moderately important, not important. Please place 5 cards in each group. Are there behaviors you think important that are unlisted? Which?

very important moderately important not important

19. These are some possible reasons for wanting to change the students' behavior. Please rank them from most important to least important from your point of view.

( ) 1. to help them feel better about themselves.
( ) 2. to help them to hold a job better.
( ) 3. to help them get along socially.
( ) 4. so that we get better discipline in school, that is, to help make the students less troublesome for the teachers and staff to manage.
( ) 5. so that the students can adjust better to the expectations of their tribes and families.

( ) 6. to help the students learn how to conform to social norms.
( ) 7. to give them practice making their own decisions so they won't be controlled by others.

20. Where did this program go right or wrong in the following areas: if wrong, how should it have been done?

1. How it was introduced to students

2. How it trained students

3. The behaviors picked to be changed

4. The method of reinforcement

5. The responsibilities of the reinforcer

6. The responsibilities of the student

7. The value of the reinforcement

8. The evaluation procedure

9. Everyone doing it their own way

10. How it was introduced to teachers

11. How it trained teachers

12. Anything else

21. How do you feel about the introduction of new educational techniques to this school?

22. Under what circumstances do you think it is appropriate for the administration to ask a teacher at the Indian School to participate in a new program?

When is it inappropriate?

"This I Believe" Instrument

Instructions

In the following pages we would like for you to write your opinions or beliefs about several topics. Since we don't want to spend too much time on this activity, we will time you on each topic at a pace that will make it necessary for you to work rapidly, but will give you enough time to record at least a couple of ideas about each referent word.

Be sure to write what you genuinely believe.

Please write on the topics in the order of their appearance. Wait to turn each page until the examiner gives you the signal, and once you have turned a page, please do not turn back to it.

When you have finished reading these instructions, wait for the signal, then turn to the first referent and begin writing.

1. This I believe about Authority.....
2. This I believe about innovation.....
3. This I believe about Friendship.....
4. This I believe about Freedom.....
5. This I believe about people who steal.....
6. This I believe about Militancy.....
7. This I believe about Education.....
8. This I believe about People on Welfare.....
9. This I believe about discipline.....
10. This I believe about people who are late.....
11. This I believe about Loyalty.....

Section VI

Final Conclusions and Recommendations

Evaluation Team

FINAL CONCLUSIONS AND RECOMMENDATIONS

After spending many weeks analyzing mountains of data, the evaluation team is left with the distinct feeling that their most important comments still must be based upon subjective rather than objective data. This is not due only to the fact that objective data is lacking in some areas, although that was a serious problem. Neither does it mean that there are no clear impressions to be gained from the analysis of data. However, no matter how extensive the body of objective data one might gather, it would still fail to convey a true impression of the project at the Albuquerque Indian School during the past year. Even more important, it would very likely convey a false impression about the prospects for the coming year, since there seem to be a number of factors that may be considerably different for the coming year. Perhaps some elaboration of this theme will help the reader to understand what is meant by such comments.

First, one should recognize that a great deal of the objective data indicates that the hopes and expectations of the project planners were not achieved to any high degree by the first year of the activity. If one considered only the impact on the total student body involved in the project, and measured objective changes in target behaviors over the period of the past few months, one would be quite disappointed to find that the shifts were not highly significant. Drinking, going AWOL, and showing general disinterest in academic work are still very common behaviors among some of the students at the school. However, before writing the project off as a failure on the basis of such findings, one must consider other very important factors.

A great many of the behaviors chosen for study and modification are very complex and very deeply rooted behaviors. Further, many of them developed because they were rewarding to the individuals, and it will take something more rewarding to replace them in the students' repertoire of behavior. Also, there are factors completely outside the control of the Albuquerque Indian School which are contributing to many of these troubles and will continue to support the undesirable behaviors in spite of things that go on at the school. Most powerful among these external factors, perhaps, is the very nature of the student body. If many of the staff members are correct, and increasing numbers of the students are sent to the school because of academic, social, or emotional difficulties, then it is too much to expect that any group project of this nature could succeed in eradicating the undesirable behaviors exhibited by such students in the brief period of one year. Indeed, it is doubtful whether such behaviors could be modified significantly by any kind of group project without extensive supplementary individual work of a rather extensive nature.

All of this does not mean, however, that the project has been a waste of time or money. Even though the group data does not indicate significant improvement in behavior, there is some evidence that changes are taking place in a number of individuals. Some children are beginning to form habits that contribute to punctuality, good attendance, and other characteristics that may be expected to help their achievement and make life easier for them in whatever academic work they may decide to pursue in the future. Many of these habits have not yet come to full fruition, and a great many students have not yet chosen to adopt such habits, but the seeds and beginnings seem to be present in a number of areas. It took years for many children to fall behind academically as far as they had, and it will take years for them to

overcome the deficits in their background and study habits to the extent that gains would be significant on a standardized test. The pattern of heavy drinking has developed because of the contribution of many interacting factors, and one can only expect that it will take the joint influence of many compensating factors to modify such behavior significantly.

In order to cope effectively with the kinds of complex and difficult problems just discussed, the staff at Albuquerque Indian School would have to present a totally unified front with maximum impact. Further, it would need to be a group of enthusiastic, highly trained, confident individuals with the best of support services at their disposal. Unfortunately, some of these factors were found to be missing by the evaluation team. Many of the staff members were not in favor of the project, and therefore did not support it at all. Others who might have provided support were not trained to the extent that they felt comfortable and confident in the use of reinforcement techniques. Still others who had both enthusiasm for the project and training in the required skills were in situations where their success was drastically limited by their circumstances. The total impact of all these factors was that the staff was nowhere near the level of effective participation and support that would be required to tackle problems of the magnitude described in the preceding section.

A third feature in this interaction of factors was the organizational lack that was found in several places. For a number of reasons that seemed primarily related to lack of time between final planning and initiation of the project, administrative organizations did not function smoothly. The introduction of the idea to the staff, the involvement of the staff in the planning phase, the collection of adequate baseline data, the provision of adequate training for staff members, the development of clear guidelines for

the carrying out of the project and provisions for the supervision of the staff in meeting the requirements of the project, all seem to have suffered from significant weaknesses during the first year. The deficiencies and problems were exacerbated because of an unfortunate lack of communication between administration and staff.

This lack of communication, in the opinions of the members of the evaluation team, cannot be attributed to lack of competence or inadequate effort on the part of either administration or staff. Both of these groups seem to be composed primarily of competent individuals who are doing their best to cope with difficult circumstances. It seems evident that the difficulty is attributable at least in part, to some basic differences in philosophy and function, together with the press of circumstances already described. A great many of the staff members are veterans of many years of service at the Albuquerque Indian School, and they have seen what they regard as a definite deterioration in the quality of the academic program and the level of morale at the school. They tend to attribute this deterioration, at least in part, to some of the recent changes and experiments by some of the young administrators. On the other hand, some of the administrators are frustrated by the lack of cooperation they receive from the staff in their attempts to cope with the increasingly difficult problems that come to the school. The result has been a communications breakdown, sometimes almost approaching a credibility gap, between administration and staff. The recommendations in several portions of this report deal specifically with the need for systematic attempts to bridge such gaps.

With all of this information available, it becomes possible to give a fairer evaluation to the data, and therefore to the project. It then becomes possible to say that, even though there were no large changes in most

of the behaviors for most of the students, there were some changes that seem highly encouraging. Some youngsters overtly changed on a number of behaviors, other children overtly changed in terms of a few behaviors, and considerable numbers of children changed with regard to one or two behaviors. More important, perhaps, is the strong probability that a great many children who have not yet overtly changed their behavior have at least covertly begun to evaluate the relative gains and losses that accrue from various kinds of behavior. They may have begun to recognize that they could, if they decided to do so, plan for and achieve some desired objective, even though that objective might be relatively small.

It also seems possible that change has begun in another area which does not show up on any of the objective information at this point. One could hope that the systematic reinforcement system is gradually replacing a system based on punishment in many of the interactions between staff members and students. A number of staff members mentioned in several instances thay they regularly give praise as a reward, but that they have some doubt about using material rewards. Therefore, even though they may not be utilizing the token rewards effectively, these staff members may be shifting toward a behavior modification system that could bear fruit in the long run. To be sure, as has been noted in previous sections of this report, it is presently important that staff members use the tokens with some consistency, in order that maximum impact may be generated. However, in the long run, it seems even more important that staff members look for good behaviors to reinforce instead of bad behaviors to punish. That seems to be happening, and one may hoep that it will continue to develop over the coming year. Along with that change, staff members should begin to recognize that the children are capable of working if they have obtainable goals and appropriate incentives. This

would be, in itself, a major breakthrough from the situation where staff members have little regard for the abilities of children to regulate their own behaviors, and little hope that the students will work to achieve their goals.

In summary, then, it would be the joint opinion of the evaluation team that the experiment has been very much worthwhile, even though objective data do not provide strong support for such an opinion. It is the opinion of the team that the goals may have been too ambitious, the problems attacked too complex, the available time too short, and the prevailing conditions too unfavorable to achieve desired results during the first year. Given the benefit of the first year of experience, together with more systematic planning, better public relations work with students and staff, improvements in the record-keeping system, rather extensive training of the staff, an incentive program for the staff, and very clear guidelines for the utilization of the program, the second year may be more successful. Even so, it seems very doubtful that one should anticipate marked changes in academic achievement or in deeply rooted complex behaviors on the part of large numbers of children who are experiencing severe difficulty in those areas. It would seem appropriate to set more realistic goals, in terms of both the numbers of children involved and the kinds of behavior to be modified. However, that is not a matter for the evaluation team to decide, but only a recommendation growing out ot this study.